

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-1

Perfect score: 119
Sequence: 1 ANDHSLLEAMSDNDPTPLMD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1475	US-09-007-999-2	Sequence 2, Appl1
2	119	100.0	1475	US-09-210-361-2	Sequence 2, Appl1
3	119	100.0	1475	US-09-740-274-2	Sequence 2, Appl1
4	110	92.4	1375	US-09-210-361-4	Sequence 4, Appl1
5	110	92.4	1375	US-09-740-274-4	Sequence 4, Appl1
6	75	63.0	545	US-09-604-957-4	Sequence 4, Appl1
7	75	63.0	1430	US-09-008-172-2	Sequence 2, Appl1
8	75	63.0	1430	US-09-210-361-6	Sequence 6, Appl1
9	75	63.0	1430	US-09-740-274-6	Sequence 6, Appl1
10	71	59.7	523	US-09-604-957-5	Sequence 5, Appl1
11	63	52.9	1577	US-08-793-824-2	Sequence 2, Appl1
12	60	50.4	584	US-09-604-957-6	Sequence 6, Appl1
13	49	41.2	2057	US-09-499-203-2	Sequence 2, Appl1
14	49	41.2	535	US-09-604-957-7	Sequence 7, Appl1
15	49	41.2	1278	US-09-604-957-3	Sequence 3, Appl1
16	46	38.7	336	US-09-198-452A-662	Sequence 662, App
17	46	38.7	596	US-09-252-991A-21255	Sequence 21255, A
18	45	37.8	236	US-09-154-750A-80	Sequence 80, Appl
19	45	37.8	385	US-08-361-920-23	Sequence 23, Appl
20	45	37.8	385	US-08-479-939-23	Sequence 23, Appl
21	45	37.8	385	US-08-483-432-23	Sequence 23, Appl
22	44.5	37.4	130	US-09-328-352-4952	Sequence 4952, Ap
23	44.5	37.4	236	US-09-252-991A-31380	Sequence 31380, A
24	44	37.0	234	US-09-134-001C-4514	Sequence 4514, Ap
25	44	37.0	384	US-09-311-170-2	Sequence 2, Appl
26	44	37.0	647	US-09-134-001C-5458	Sequence 5458, Ap
27	44	37.0	749	US-09-562-737-96	Sequence 96, Appl

28	43	36.1	2465	2	US-08-596-291-3	Sequence 3, Appl1
29	43	36.1	2465	3	US-09-100-804-3	Sequence 3, Appl1
30	43	36.1	2466	3	US-09-080-855-12	Sequence 12, Appl
31	43	36.1	2466	4	US-09-566-076-12	Sequence 12, Appl
32	43	36.1	2466	5	PCT-US94-09943-2	Sequence 2, Appl1
33	43	36.1	2485	3	US-09-290-640-46	Sequence 46, Appl
34	43	35.3	211	4	US-09-107-532A-5029	Sequence 5029, Ap
35	42	35.3	259	1	US-08-277-231A-3	Sequence 3, Appl1
36	42	35.3	259	2	US-08-473-750-6	Sequence 6, Appl1
37	42	35.3	259	2	US-08-477-326-6	Sequence 6, Appl1
38	42	35.3	288	4	US-09-252-991A-23168	Sequence 23168, A
39	42	35.3	989	2	US-08-070-301-16	Sequence 16, Appl
40	42	35.3	1428	4	US-09-252-991A-30731	Sequence 30731, A
41	41.5	34.9	454	4	US-09-252-991A-28000	Sequence 28000, A
42	41.5	34.9	1048	4	US-09-171-699-10	Sequence 10, Appl
43	41	34.5	244	4	US-09-252-991A-22373	Sequence 22373, A
44	41	34.5	270	4	US-09-323-872A-38	Sequence 38, Appl
45	41	34.5	296	4	US-09-328-352-7769	Sequence 7769, Ap

ALIGNMENTS

```
RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007, 999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478, 704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-007-999-2

Query Match      100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANDHSLLEAMSDNDPTPLMD 21
Db      481 ANDHSLLEAMSDNDPTPLMD 501

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210, 361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007, 999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478, 704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009, 620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485, 243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008, 172
```

```
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match          100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLHD 21
Db 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match          100.0%; Score 119; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLHD 21
Db 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match          92.4%; Score 110; DB 3; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLHD 21
Db 507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          92.4%; Score 110; DB 4; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLHD 21
Db 507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
```

```
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRF
; ORGANISM: Streptococcus mutans
; US-09-604-957-4
```

```
Query Match          63.0%; Score 75; DB 4; Length 545;
Best Local Similarity 71.4%; Pred. No. 0.00043;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAWSDNDTPYLHD 21
Db 75 AINHLSTLEAWSDNDPQYKND 95
```

```
RESULT 7
; US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: Streptococcus mutans
; US-09-008-172-2
```

```
Query Match          63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAWSDNDTPYLHD 21
Db 495 AINHLSTLEAWSDNDPQYKND 515
```

```
RESULT 8
; US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
```

```
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: streptococcus mutans
; US-09-210-361-6
```

```
Query Match          63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAWSDNDTPYLHD 21
Db 495 AINHLSTLEAWSDNDPQYKND 515
```

```
RESULT 9
; US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: streptococcus mutans
; US-09-740-274-6
```

```
Query Match          63.0%; Score 75; DB 4; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAWSDNDTPYLHD 21
Db 495 AINHLSTLEAWSDNDPQYKND 515
```

```
RESULT 10
; US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6466314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
```

APPLICANT: RAHAOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 5
LENGTH: 523
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 59.7%; Score 71; DB 4; Length 523;
Best Local Similarity 66.7%; Pred. No. 0.0018;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANDHSLTEAWSDNDTPYLHD 21
Db 75 ANQHSLTEAWSDNDTPYLHD 95

RESULT 11
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of Plants to
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Griffiths Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.14;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HSLTEAWSDNDTPYLHD 21
Db 591 HSLTEAWSDNDTPYLHD 608

RESULT 12
US-09-604-957-6
Sequence 6, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHOTTEN, GERRITDINA HENDRIKA
APPLICANT: DIKHUIZEN, LUBBERT
APPLICANT: RAHAOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 584
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 50.4%; Score 60; DB 4; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.13;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANDHSLTEAWSDNDTPYLHD 20
Db 75 ANKHSLTEAWSDNDTPYLHD 94

RESULT 13
US-09-499-203-2
Sequence 2, Application US/09499203
Patent No. 6570065
GENERAL INFORMATION:
APPLICANT: KOSSMANN, Jens
APPLICANT: WELSH, Thomas
APPLICANT: QUANZ, Martin
APPLICANT: KNUTH, Karola
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
FILE REFERENCE: 147-196P
CURRENT APPLICATION NUMBER: US/09/499,203
CURRENT FILING DATE: 2000-02-08
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 2057
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 50.4%; Score 60; DB 4; Length 2057;
Best Local Similarity 55.0%; Pred. No. 0.57;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANDHSLTEAWSDNDTPYLHD 20
Db 665 ANKHSLTEAWSDNDTPYLHD 684

RESULT 14
US-09-604-957-7
Sequence 7, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:


```

?
? APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
? APPLICANT: DIKHUIZEN, LUBBERT
? APPLICANT: RAHOUI, HAKIM
? APPLICANT: LEER, ROBERT-JAN
? TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
? FILE REFERENCE: BO 43388
? CURRENT APPLICATION NUMBER: US/09/604,957
? CURRENT FILING DATE: 2000-06-28
? PRIOR APPLICATION NUMBER: 00201871.1
? PRIOR FILING DATE: 2000-05-25
? NUMBER OF SEQ ID NOS: 17
? SOFTWARE: Patent Ver. 2.1
? SEQ ID NO 7
? LENGTH: 535
? TYPE: PR1
? ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

Query Match	41.2%;	Score 49;	DB 4;	Length 535;
Best Local Similarity	40.0%;	Pred. No. 7;		
Matches	8;	Conservative	5;	Mismatches 7; Indels 0; Gaps 0;

```
QY      1 ANDHLSILEAWSDNDTPYLH 20
        :|/::|||/|:|:|:
Db      74 SNKHINILLEDMNHADPEYFN 93
```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUTAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PR1
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

```

Query Match	41.2%	Score 49	DB 4	Length 1278
Best Local Similarity	40.0%	Pred. NO. 20		
Matches 8	Conservative 5	Mismatches 7	Indels 0	Gaps 0

QY 1 ANDHLSILEAWSNDNTPYLH 20
:|::|||:|:
Db 550 SNKHINILEDWNHADPEYFN 569

Search completed: November 13, 2003, 09:47:57
Job time : 15.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHLSILFAWSNDPTLYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1017	23	AAU79285
2	119	100.0	1475	23	AAU98027
3	119	100.0	1475	23	AAU98030
4	119	100.0	1475	23	AAU98031
5	119	100.0	1475	23	AAU98032
6	119	100.0	1475	23	AAU98033
7	119	100.0	1475	23	AAU98034
8	119	100.0	1475	23	AAU98035
9	119	100.0	1475	23	AAU98036

10	119	100.0	1475	23	AAU98037	S. mutans glucosyl
11	119	100.0	1475	23	AAU98038	S. mutans GTFB mut
12	119	100.0	1475	23	AAU98039	S. mutans glucosyl
13	119	100.0	1475	23	AAU98040	S. mutans glucosyl
14	119	100.0	1475	23	AAU98041	Streptococcus muta
15	119	92.4	1375	23	AAU98028	S. mutans glucosyl
16	110	92.4	1375	23	AAU98029	Streptococcus muta
17	108	90.8	1375	23	AAU98030	Streptococcus muta
18	75	63.0	1430	23	AAU98041	Glucosyltransferase
19	75	63.0	1430	23	AAU98042	S. mutans glucosyl
20	75	63.0	1430	23	AAU98043	S. mutans glucosyl
21	75	63.0	1430	23	AAU98044	S. mutans glucosyl
22	75	63.0	1430	23	AAU98045	S. mutans glucosyl
23	75	63.0	1430	23	AAU98046	S. mutans glucosyl
24	71	59.7	1527	23	AAU98005	Leuconostoc mesent
25	69	58.0	2835	23	ABB98574	Dextran saccharase
26	66	55.5	12	23	ABB98641	Dextrane-saccharas
27	63	52.9	1577	17	AAU91047	Alpha-D-glucosyltr
28	60	50.4	12	23	ABB98642	Dextrane-saccharas
29	59	49.6	12	23	ABB98643	Dextrane-saccharas
30	59	49.6	12	23	ABB98644	Dextrane-saccharas
31	53	44.5	12	23	ABB98645	Dextrane-saccharas
32	53	44.5	12	23	ABB98646	Dextrane-saccharas
33	51	42.9	486	21	AAU36552	Lawsonia intracell
34	50	42.0	305	21	AAU60945	Arabidopsis thalia
35	50	42.0	339	21	AAU60944	Arabidopsis thalia
36	50	42.0	649	22	AAU62860	Prothiobacterium
37	49	41.2	12	23	ABB98578	Dextrane-saccharas
38	49	41.2	12	23	ABB98648	Dextrane-saccharas
39	49	41.2	302	22	AAU38411	Salmonella typhi c
40	49	41.2	1781	23	AAU74519	Lactobacillus reut
41	48.5	40.8	401	22	AAU43544	Human polypeptide
42	48	40.3	855	23	ABB98573	Dextran saccharase
43	47.5	39.9	534	21	AAU67414	Arabidopsis aldehy
44	47	39.5	195	21	AAU60941	Arabidopsis thalia
45	47	39.5	338	15	AAU65965	T. niivum GAPDH.

ALIGNMENTS

RESULT 1
AAU79285
AAU79285 standard; Protein: 1017 AA.
AC AAU79285;
DT 13-AUG-2002 (first entry)
XX
DE Streptococcus mutans monoclonal antibody-related protein #2.
XX
KW Antibody; dental caries; water insoluble glucan synthetase;
XX anti-caries; glucosyl transferase-B; immunotherapy.
XX
OS Streptococcus mutans.
XX
PN JP2002114709-A.
XX
PD 16-APR-2002.
XX
PF 04-OCT-2000; 2000JP-0304889.
XX
PR 04-OCT-2000; 2000JP-0304889.
XX
PA (UN1-) UNIV NIPPON.
XX
XX WPI; 2002-448101/48.
XX
PT Anti-caries agent composed of a monoclonal antibody against an
PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT transferase-B (GTF-B) of Streptococcus mutans -
XX
PS Claim 4; Page 17-19; 28pp; Japanese.

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-caries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 XX Sequence 1017 AA;
 SQ
 Query Match 100.0%; Score 119; DB 23; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 3,5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ANDHLSILEAWSNDNTPYLHD 21
 DB 447 ANDHLSILEAWSNDNTPYLHD 467
 RESULT 2
 AAU98027 standard; Protein; 1475 AA.
 ID AAU98027;
 AC AAU98027;
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture.
 XX
 XX Streptococcus mutans.
 OS
 XX US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 PR
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 XX WPI; 2002-414332/44.
 DR N-PDB; ABK52938.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Disclosure; Page 21-25; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant, growing the plant cell under plant growing
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 CC
 XX Sequence 1475 AA;
 SQ
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ANDHLSILEAWSNDNTPYLHD 21
 DB 481 ANDHLSILEAWSNDNTPYLHD 501
 RESULT 3
 AAU98030 standard; Protein; 1475 AA.
 ID AAU98030
 AC AAU98030;
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB mutant 1448V.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 XX Streptococcus mutans.
 OS
 XX Synthetic.
 PA
 XX Key Location/Qualifiers
 FH MISC-difference 448 /note= "Wild-type Ile substituted by Val"
 FT
 XX US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query March 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. NO. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHSLILEAWSNDPTPIHLD 21
 Db 481 ANDHSLILEAWSNDPTPIHLD 501
 RESULT 4
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX
 DT 27-AUG-2002 (first entry)

XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Msc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 XX
 PN US2002031826-A1.
 PD
 XX 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and

```

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
CC
XX
SQ Sequence 1475 AA;
Oy
Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5,4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps
1
1 ANDHLSTLEAMSDNDPTPLHD 21
|||||
481 ANDHLSTLEAMSDNDPTPLHD 501
RESULT 5
ID AAU98032
AC AAU98032;
DT 27-AUG-2002 (first entry)
DE S. mutans glucosyltransferase GTFB mutant D567T.
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutain.
OS Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX
PI Nichols SE;
XX
DR WPI, 2002-414332/44.
XX
PT Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
PS Claim 36; Page -, 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169N/Y170M/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from 7589D, 7589E, N471D, N471D/7589D, and
CC N471D/7589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary

```

CC	polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
CC	an expression cassette comprising the polynucleotide operably linked to a
CC	promoter, a vector comprising the expression cassette, host cell
CC	introduced with the vector, a transgenic plant comprising the
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC	coating composition comprising a glucan produced in a plant transformed
CC	with a gene encoding the mutant GFP, wild type or, starch, a latex,
CC	thermoplastic molecule or their combinations or glucan and starch where
CC	the glucan is produced in the amyloplast and/or vacuole of a maize line
CC	deficient in starch biosynthesis, transformed with a gene encoding a
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC	comprising the glucan (paper sizing/coating agent). The vector is useful
CC	for producing a glucan in a plant. The method comprises transforming a
CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transit sequence from
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GFP, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GFP mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GFPB sequence appearing as AAU98027
CC	and the information in claim 36.
XX	
SQ	Sequence 1475 AA;
XX	
Query Match	100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity	100.0%; Pred. No. 5,4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0,	
QY	1 ANDHSLTEAWSDNDPTYLHD 21
DB	481 ANDHSLTEAWSDNDPTYLHD 501
RESULT 6	
AAU98033	
XX	AAU98033 standard; Protein; 1475 AA.
XX	
AC	AAU98033;
XX	
DT	27-ADG-2002 (first entry)
XX	
XX	S. mutans glucosyltransferase GFPB mutant K1014T.
XX	
XX	Glucosyltransferase; GFPB; transgenic plant; paper sizing;
KW	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amyloplast; vacuole; paper manufacture; mutant; mutein.
XX	
OS	Streptococcus mutans.
XX	
XX	Synthetic.
XX	
XX	Key Location/Qualifiers
FT	Misc-difference 1014
FT	/note= "Wild-type Lys substituted by Thr"
PN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
XX	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHSLTEAMSDNTPRYLHD 21
 DB 481 ANDHSLTEAMSDNTPRYLHD 501
 RESULT 7
 AAU98034
 ID AAU98034 standard; Protein; 1475 AA.
 XX
 AC AAU98034;

XX
 DT 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N/D567T.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX Glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSILEAWSNDPTPLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTPLHD 501

RESULT 9
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.

XX AAU98035;

XX 27-AUG-2002 (first entry)

DE 5. mutans glucosyltransferase GTFB mutant D457N/D571K.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FT US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful

CC as substitutes for and additions to modified starch and latexes in
 CC paper manufacture, comprises mutations in specific positions
 CC Claim 36; Page -; 44p; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K.

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K7799/K1014T,

CC Y169A/Y10A/Y171A, and K779Q or a GTF D polypeptide having

CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the

CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

CC coating composition comprising a glucan produced in a plant transformed

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where

CC the glucan is produced in the amyloplast and/or vacuole or a maize line

CC deficient in starch biosynthesis, transformed with a gene encoding a

CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful

CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing

CC conditions to produce a regenerated plant and inducing expression of the

CC polynucleotide for a time sufficient to produce the glucan in the

CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC is produced in the amyloplast of potato or and additions to modified starch

CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper

CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.

CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSILEAWSNDPTPLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTPLHD 501

RESULT 9
 AAU98036
 ID AAU98036 standard; Protein; 1475 AA.

XX AAU98036;

XX 27-AUG-2002 (first entry)

DE 5. mutans glucosyltransferase GTFB mutant D567T/D571K.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FT US2002031826-A1.

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes p1 or p2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLHD 21
 |||||

DB 481 ANDHSLTEAMSNDTPYLHD 501
 XX
 XX RESULT 10
 XX AAU98037
 XX ID AAU98037 standard; Protein; 1475 AA.
 XX
 XX AC AAU98037;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
 XX
 XX Glucosyltransferase; GTFB; transgenic plant; paper string;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muteln.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"
 FT /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014
 FT /note= "Wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes p1 or p2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GFP mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

XX Query Match 100.0%; Score 119; DB 23; Length 1475;
 XX Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLHD 21
 Db 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 11
 AAU98038 ID AAU98038 standard; Protein; 1475 AA.

XX AC AAU98038;

XX DT 27-AUG-2002 (first entry)

DE S. mutans GFPB mutant 1448V/D457N/D567T/D571K/K779Q/K1014T.

XX Glucosyltransferase; GFPB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.

OS Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"

FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"

FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX Claim 36, Page -; 44pp: English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
 CC a thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GFP mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

XX Query Match 100.0%; Score 119; DB 23; Length 1475;
 XX Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLHD 21
 Db 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 12

AAU98039 ID AAU98039 standard; Protein; 1475 AA.

XX AC AAU98039;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.

DE Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

PH Msc-difference 169..171

FT /note= "Wild-type Tyr-Tyr-Tyr substituted by

FT Ala-Ala-Ala"

XX US2002031826-A1.

PN 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

PA Nichols SE;

XX WPI; 2002-414332/44.

DR Glucosyltransferase B or D protein useful for producing a glucan useful

PT as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -; 44pp; English.

PS The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from 1448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,

CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having

CC changes at positions from T589D, T589E, N471D, N471D/T589D and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the

CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

CC coating composition comprising a glucan produced in a plant transformed

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where

CC the glucan is produced in the amyloplast and/or vacuole or a maize line

CC deficient in starch biosynthesis, transformed with a gene encoding a

CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful

CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing

CC conditions to produce a regenerated plant and inducing expression of the

CC polynucleotide for a time sufficient to produce the glucan in the

CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC is produced in the amyloplast of potato or the vacuole of sugar beet.

CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper

CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.

CC Note: The present sequence is not shown in the specification but

CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.

XX SQ Sequence 1475 AA;

XX Query Match 100.0%; Score 119; DB 23; Length 1475;

XX Best Local Similarity 100.0%; Pred. No. 5.4e-09;

XX Matches 21; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDPTPYLHD 21

DB 481 ANDHSLILEAWSNDPTPYLHD 501

RESULT 13

AAU98040 ID AAU98040 standard; Protein; 1475 AA.

AAU98040; AC AAU98040;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant K779Q.

DE Glucosyltransferase; GTFB; transgenic plant; paper sizing;

XX coating composition; glucan; starch; latex; thermoplastic molecule;

KW amyloplast; vacuole; paper manufacture; mutant; mutein.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

PH Msc-difference 779

FT /note= "Wild-type Lys substituted by Gln"

FT US2002031826-A1.

PN 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

PA Nichols SE;

XX WPI; 2002-414332/44.

DR Glucosyltransferase B or D protein useful for producing a glucan useful

PT as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -; 44pp; English.

PS The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from 1448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,

CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having

CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T569E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll A3 binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTYLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTYLHD 501

RESULT 14

AAU79284 standard; Protein; 1476 AA.

AAU79284;

13-AUG-2002 (first entry)

Streptococcus mutans monoclonal antibody-related protein #1.

Antibody; dental caries; water insoluble glucan synthetase;

anti-carries; glucosyl transferase-B; immunotherapy.

Streptococcus mutans.

JP2002114709-A.

16-APR-2002.

04-OCT-2000; 2000JP-0304889.

04-OCT-2000; 2000JP-0304889.

(UNIT-) UNIT NIPPON.

WPI; 2002-448101/48.

Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans
 PS Claim 3; Page 13-16; 28pp; Japanese.

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.

Sequence 1476 AA;

Query Match 100.0%; Score 119; DB 23; Length 1476;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTYLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTYLHD 501

RESULT 15

AAU98028 standard; Protein; 1375 AA.

AAU98028;

27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;

coating composition; glucan; starch; latex; thermoplastic molecule;

amyloplast; vacuole; paper manufacture.

Streptococcus mutans.

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

11-DEC-1998; 98US-0210361.

07-JUN-1995; 95US-0478704.

07-JUN-1995; 95US-0482711.

07-JUN-1995; 95US-0485243.

16-JAN-1998; 98US-0007999.

16-JAN-1998; 98US-0008172.

20-JAN-1998; 98US-0009620.

(NICH/) NICHOLS S E.

Nichols SE;

WPI; 2002-414332/44.

N-PSDB; ABR52939.

Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions
 Disclosure; Page 30-33; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position 1448V, D457N,
 D567T, K1014I, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GFPc.
 CC XX

SQ Sequence 1375 AA;

Query Match 92.4%; Score 110; DB 23; Length 1375;

Best Local Similarity 95.2%; Pred. No. 1.1e-07;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTPYLHD 21
 |||||
 DB 507 ANDHLSILEAWSNDPTPYLHD 527

Search completed: November 13, 2003, 09:38:25
 Job time : 39.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHSLIEMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues
Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubppaa/PCOT_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppaa/PCOTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1475	9	US-09-740-274-2
2	110	92.4	1375	9	US-09-740-274-4
3	75	63.0	545	10	US-09-995-749A-10
4	75	63.0	1430	9	US-09-740-274-6
5	71	59.7	522	10	US-09-995-749A-11
6	60	50.4	584	10	US-09-995-749A-12
7	51	42.9	302	12	US-10-009-823A-1
8	49	41.2	302	9	US-09-815-242-14004
9	49	41.2	535	10	US-09-995-749A-13
10	49	41.2	1781	10	US-09-995-749A-2
11	47.5	39.9	534	10	US-09-344-882-24
12	47.5	39.9	534	15	US-10-293-865-24
13	47	39.5	15	12	US-10-295-693-90
14	47	39.5	15	15	US-10-067-649-90
15	46.5	39.1	1604	9	US-09-888-615-73

16	46	38.7	199	10	US-09-738-626-4571	Sequence 4571, Ap
17	46	38.7	476	11	US-09-908-299-9	Sequence 34, Appl
18	46	38.7	539	11	US-10-230-026-34	Sequence 94, Appl
19	45.5	38.2	156	12	US-10-001-245-125	Sequence 125, App
20	45	37.8	268	15	US-10-156-761-12823	Sequence 12823, A
21	44.5	37.4	210	15	US-10-127-032-167	Sequence 167, App
22	44	37.0	322	9	US-09-815-242-11634	Sequence 9158, Ap
23	44	37.0	331	9	US-09-815-242-10188	Sequence 11634, A
24	44	37.0	334	9	US-09-815-242-10188	Sequence 10188, A
25	44	37.0	749	15	US-10-221-962-96	Sequence 13981, A
26	44	37.0	816	15	US-10-080-114A-7	Sequence 96, Appl
27	43.5	36.6	155	12	US-10-001-245-126	Sequence 126, App
28	43	36.1	41	11	US-09-764-891-3308	Sequence 3308, Ap
29	43	36.1	311	11	US-09-878-781-18	Sequence 18, Appl
30	43	36.1	381	11	US-09-095-478-9	Sequence 9, Appl
31	43	36.1	381	11	US-09-095-478-9	Sequence 7, Appl
32	43	36.1	715	15	US-10-156-761-7562	Sequence 7562, Ap
33	43	36.1	1267	15	US-10-059-585-56	Sequence 56, Appl
34	43	36.1	2466	12	US-10-177-980-12	Sequence 12, Appl
35	43	36.1	2485	9	US-09-802-669-46	Sequence 46, Appl
36	43	36.1	2654	15	US-10-227-610-2	Sequence 2, Appl
37	43	36.1	491	9	US-09-815-242-10190	Sequence 10190, A
38	42.5	35.7	491	9	US-09-815-242-10190	Sequence 13756, A
39	42.5	35.7	31	15	US-10-092-908-40	Sequence 40, Appl
40	42	35.3	41	15	US-10-092-908-41	Sequence 41, Appl
41	42	35.3	95	9	US-09-945-301-14	Sequence 14, Appl
42	42	35.3	298	11	US-09-988-067B-26	Sequence 26, Appl
43	42	35.3	348	10	US-09-738-626-6855	Sequence 6855, Ap
44	42	35.3	484	15	US-10-156-761-8339	Sequence 8339, Ap
45	42	35.3				

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FaetsEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 100.0%; Score 119; DB 9; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ANDHSLIEMSDNDTPYLHD 21

Db 481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 92.4%; Score 110; DB 9; Length 1375;

Best Local Similarity 95.2%; Pred. No. 4e-08; 1; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 21

Db 507 ANDHLSILEAMSDNDTPYLHD 527

RESULT 3

US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOUBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 63.0%; Score 75; DB 10; Length 545;

Best Local Similarity 71.4%; Pred. No. 0.0033; 5; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 21
Db 75 AINHLSILEAMSDNDPQYNKD 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 63.0%; Score 75; DB 9; Length 1430;

Best Local Similarity 71.4%; Pred. No. 0.0097; 5; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 21

Db 495 AINHLSILEAMSDNDPQYNKD 515

RESULT 5

US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOUBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 59.7%; Score 71; DB 10; Length 522;

Best Local Similarity 66.7%; Pred. No. 0.013;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSLEAMSNDTPYLD 21
Db 75 ANQHLSTLEDSHNDPLXYTD 95

RESULT 6

US-09-995-749A-12
Sequence 12, Application US/09995749A
Patent No. US20020155568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUITZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 584
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 50.4%; Score 60; DB 10; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.72;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSLEAMSNDTPYLD 20
Db 75 ANKHLSTLEDSHNDPLXYTD 94

RESULT 7

US-10-009-823A-1
Sequence 1, Application US/10009823A
Publication No. US20030157120A1
GENERAL INFORMATION:
APPLICANT: Panaccio, Michael
APPLICANT: Rosey, Everett Lee
APPLICANT: Sinistraj, Meri
APPLICANT: Haase, Detlef
APPLICANT: Parsons, Jim
APPLICANT: Ankenbauer, Robert G.
TITLE OF INVENTION: LAWSONIA DERIVED GENE AND RELATED FLGE
FILE REFERENCE: DAVI150.001APC
CURRENT APPLICATION NUMBER: US/10/009,823A
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: PCT/AU00/00437
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: US 60/133,973
PRIOR FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 502
TYPE: PRT
ORGANISM: Lawsonia intracellularis
US-10-009-823A-1

Query Match 42.9%; Score 51; DB 12; Length 502;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSLEAMSNDTPYLD 17
Db 185 ANPYFALLESMKNGTNP 201

RESULT 8

US-09-815-242-14004
Sequence 14004, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: BLITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14004
LENGTH: 302
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-14004

Query Match 41.2%; Score 49; DB 9; Length 302;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSLEAMSNDTPYLD 21
Db 267 NDRVLIVDEVLEASDIPYRHD 266

RESULT 9

US-09-995-749A-13
Sequence 13, Application US/09995749A
Patent No. US20020155568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUITZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13
LENGTH: 535
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 41.2%; Score 49; DB 10; Length 535;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHLSLEAWSNDTPYLH 20
DB 74 SNKHINILEDWVHADPEYFN 93

RESULT 10
US-09-995-749A-2
Sequence 2, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA, HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: BO43388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1781
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 41.2%; Score 49; DB 10; Length 1781;
Best Local Similarity 40.0%; Pred. No. 1,2e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHLSLEAWSNDTPYLH 20
DB 1053 SNKHINILEDWVHADPEYFN 1072

RESULT 11
US-09-344-882-24
Sequence 24, Application US/09344882
Patent No. US20020162137A1
GENERAL INFORMATION:
APPLICANT: Nikolau, Basil J
APPLICANT: Wurttele, Eve S
APPLICANT: Oliver, David J
APPLICANT: Behal, Robert
APPLICANT: Schnable, Patrick S
APPLICANT: Ke, Jinsan
APPLICANT: Johnson, Jerry L
APPLICANT: Allred, Carolyn C
APPLICANT: Fatland, Beth
APPLICANT: Lutziger, Isabelle
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
TITLE OF INVENTION: Acetyl CoA Levels in Plants
FILE REFERENCE: 201573
CURRENT APPLICATION NUMBER: US/09/344,882
CURRENT FILING DATE: 1999-06-25
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn Ver. 2.2
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-344-882-24

Query Match 39.9%; Score 47.5; DB 10; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHLSLEAWSNDTPY 18
DB 132 NDHLSLEAWSNDTPY 147

RESULT 12
US-10-293-865-24
Sequence 24, Application US/10293865
Publication No. US20030106090A1
GENERAL INFORMATION:
APPLICANT: Nikolau, Basil J
APPLICANT: Wurttele, Eve S
APPLICANT: Oliver, David J
APPLICANT: Behal, Robert
APPLICANT: Schnable, Patrick S
APPLICANT: Ke, Jinsan
APPLICANT: Johnson, Jerry L
APPLICANT: Allred, Carolyn C
APPLICANT: Fatland, Beth
APPLICANT: Lutziger, Isabelle
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
TITLE OF INVENTION: Acetyl CoA Levels in Plants
FILE REFERENCE: 217113
CURRENT APPLICATION NUMBER: US/10/293,865
CURRENT FILING DATE: 2002-11-13
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/344,882
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-293-865-24

Query Match 39.9%; Score 47.5; DB 15; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHLSLEAWSNDTPY 18
DB 132 NDHLSLEAWSNDTPY 147

RESULT 13
US-10-295-693-90
Sequence 90, Application US/10295693
Publication No. US20030198976A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGRPMW14, RELATED TO THE
TITLE OF INVENTION: GPCR, GPCR3
FILE REFERENCE: D0118 CIP
CURRENT APPLICATION NUMBER: US/10/295,693
CURRENT FILING DATE: 2002-11-14
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/266,525
PRIOR FILING DATE: 2002-02-05

PRIOR APPLICATION NUMBER: US 60/329,897
 PRIOR FILING DATE: 2001-10-16
 NUMBER OF SEQ ID NOS: 95
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 90
 LENGTH: 15
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthesized Random Peptide.
 US-10-295-693-90

Query Match 39.5%; Score 47; DB 12; Length 15;
 Best Local Similarity 61.5%; Pred. No. 1.2; 2; Indels 0; Gaps 0;
 Matches 8; Conservative 3; Mismatches 0;

QY 8 LEAMSDNDTPYLH 20
 Db 3 LEAMDLSDTPHLY 15

RESULT 14
 US-10-067-649-90
 Sequence 90, Application US/10067649
 Publication No. US20030100057A1
 GENERAL INFORMATION:
 APPLICANT: Bristol-Myers Squibb Company
 TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HSPRMY14, RELATED TO
 TITLE OF INVENTION: ORPHAN GPCR, GPR73
 FILE REFERENCE: D0118 NP
 CURRENT APPLICATION NUMBER: US/10/067,649
 CURRENT FILING DATE: 2002-02-05
 PRIOR APPLICATION NUMBER: US 60/266,525
 PRIOR FILING DATE: 2001-02-05
 PRIOR APPLICATION NUMBER: US 60/329,897
 PRIOR FILING DATE: 2001-10-16
 NUMBER OF SEQ ID NOS: 92
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 90
 LENGTH: 15
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthesized Random Peptide.
 US-10-067-649-90

Query March 39.5%; Score 47; DB 15; Length 15;
 Best Local Similarity 61.5%; Pred. No. 1.2; 2; Indels 0; Gaps 0;
 Matches 8; Conservative 3; Mismatches 0;

QY 8 LEAMSDNDTPYLH 20
 Db 3 LEAMDLSDTPHLY 15

RESULT 15
 US-09-888-615-73
 Sequence 73, Application US/09888615
 Patent No. US20020064856A1
 GENERAL INFORMATION:
 APPLICANT: PLOWMAN, GREGORY
 APPLICANT: WHITE, DAVID
 APPLICANT: CAENEPEEL, SEAN
 APPLICANT: CHARYDOZAK, GLEN
 APPLICANT: MANNING, GERARD
 APPLICANT: SUDARSANAM, SUCHA
 TITLE OF INVENTION: NOVEL PROTEASES
 FILE REFERENCE: 038602/1214
 CURRENT APPLICATION NUMBER: US/09/888,615
 CURRENT FILING DATE: 2001-06-26
 PRIOR APPLICATION NUMBER: 60/214,047
 PRIOR FILING DATE: 2000-06-26
 NUMBER OF SEQ ID NOS: 150

SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 73
 LENGTH: 1604
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-888-615-73

Query Match 39.1%; Score 46.5; DB 9; Length 1604;
 Best Local Similarity 47.4%; Pred. No. 2.6e+02;
 Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 5 LSLLEAMSDN--DTPYLH 20
 Db 294 VALLEWVKDNRRTDDIPHLH 312

Search completed: November 13, 2003, 10:28:59
 Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHLSTLEAWSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1476	1	GTPB_STRMU P08987 streptococc
2	110	92.4	1455	1	GTFP_STRMU P13470 streptococc
3	108	90.8	1592	1	GTF2_STRDO P27470 streptococc
4	108	90.8	1597	1	GTF1_STRDO P11001 streptococc
5	78	65.5	1365	1	GTF5_STRDO P29336 streptococc
6	75	63.0	1462	1	GTFD_STRMU P49331 streptococc
7	57	47.9	330	1	GTPC_LEIME P01558 leishmania
8	49	41.2	302	1	YBHK_SALTI P58586 salmonella
9	49	41.2	302	1	YBHK_SALTY P58586 salmonella
10	49	41.2	336	1	UL16_EBV P03221 Epstein-Barr
11	47.5	39.9	330	1	G3PC_TRYBB P10097 trypanosoma
12	47	39.5	337	1	G3P_COCP0 P43113 coccidioidide
13	46	38.7	275	1	SC35_YEAST P53271 saccharomyc
14	46	38.7	335	1	G3P_CHUPN P09270 chlamydia p
15	46	38.7	476	1	PCHA_PSEAE P01508 pseudomonas
16	45	37.8	236	1	GAMT_HUMAN Q44353 homo sapien
17	45	37.8	385	1	GAMT_FUSOX P46239 fusarium ox
18	44.5	37.4	470	1	AMP2_STRDO P06394 streptomyc
19	44	37.0	234	1	G3P_MEIGA P05672 escherichia
20	44	37.0	330	1	G3PI_ECOLI P24165 salmonella
21	44	37.0	330	1	G3PI_SALTY P00356 gallus gall
22	44	37.0	332	1	G3P_CHICK P04379 columba liv
23	44	37.0	332	1	G3P_COLLI P08735 zea mays (m
24	44	37.0	337	1	G3PD_MAIZE P009054 hordeum vul
25	44	37.0	337	1	G3PX_HORVU P26511 hordeum vul
26	44	37.0	337	1	G3P_CRYPA P18089 cyphonectr
27	44	37.0	337	1	G3P_CRYPA P18089 cyphonectr
28	44	37.0	363	1	G3PI_JACOR P49036 jaculus ori
29	44	37.0	816	1	SUS2_MAIZE P41540 oryctolagus
30	43.5	36.6	115	1	TKN1_RABIT P060541 myocoeleacu
31	43.5	36.6	130	1	TKN1_MESAT P41539 mus musculu
32	43.5	36.6	130	1	TKN1_MOUSE P41539 mus musculu
33	43.5	36.6	130	1	TKN1_RAT P06767 rattus norv

34	43.5	36.6	491	1	G6PD_ERMCH P37986 erwina chr
35	43	36.1	174	1	PADC_VIRCH O9KRX2 vibrio chol
36	43	36.1	235	1	GAMT_RAT P10868 rattus norv
37	43	36.1	236	1	GAMT_MOUSE O35969 mus musculu
38	43	36.1	320	1	G3P_BOVIN P10096 bos taurus
39	43	36.1	322	1	G3P_SHEEP O28554 ovis aries
40	43	36.1	332	1	G3P_PIG P00355 sus scrofa
41	43	36.1	332	1	G3P_RABIT P46406 oryctolagus
42	43	36.1	334	1	G3P2_HUMAN P04406 homo sapien
43	43	36.1	337	1	G3P_AJECA O9HFX1 ajellomyces
44	43	36.1	337	1	G3P_MYCGB P47543 mycoplasma
45	43	36.1	424	1	IYD_RAT P12007 rattus norv

ALIGNMENTS

RESULT 1
GTFB_STRMU STANDARD; PRT; 1476 AA.
ID GTFB_STRMU 069381; 069384; 069387; 069390; 069396;
AC P08987; 069381; 069384; 069387; 069390; 069396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE (Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (dextranucrase)
GN GTFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxId=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=DA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans DA159, a cariogenic dental
pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6'-alpha-D-glucosyl)) (N) = D-
fructose + ((1,6)-alpha-D-glucosyl)) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGS). GTF-S SYNTHESIZES BOTH
WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; M17361; AAA89588.1; -
DR EMBL; D88651; BAA26101.1; -
DR EMBL; D88654; BAA26105.1; -
DR EMBL; D88657; BAA26109.1; -
DR EMBL; D88660; BAA26113.1; -
DR EMBL; D88677; BAA26119.1; -
DR EMBL; AE014940; AAN58705.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF00324; Glyco_hydro_70; 1.
DR Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW Complete proteome.
KW SIGNAL
FT CHAIN 1 34
FT DOMAIN 35 1476
FT DOMAIN 1097 1476
FT REPEAT 1097 1130
FT DOMAIN 1161 1470
FT REPEAT 1161 1210
FT REPEAT 1225 1275
FT REPEAT 1290 1340
FT REPEAT 1355 1405
FT REPEAT 1420 1470
FT REPEAT 62 62
FT VARIANT 65 65
FT VARIANT 68 68
FT VARIANT 78 78
FT VARIANT 86 86
FT VARIANT 89 89
FT VARIANT 168 168
FT VARIANT 276 276
FT VARIANT 399 399
FT VARIANT 474 474
FT VARIANT 512 512
FT VARIANT 519 519
FT VARIANT 701 701
FT VARIANT 708 708
FT VARIANT 938 938
FT VARIANT 952 952
FT VARIANT 963 964
FT VARIANT 968 970
FT VARIANT 1086 1086
FT VARIANT 1158 1158
FT VARIANT 1163 1163
FT VARIANT 1168 1168
FT VARIANT 1182 1182
FT VARIANT 1234 1234
FT VARIANT 1263 1263
FT VARIANT 1264 1264
FT VARIANT 1272 1272
FT VARIANT 1329 1329
FT VARIANT 1394 1394
FT VARIANT 1402 1402

POTENTIAL.
GLUCOSYLTRANSFERASE-1.
CATALYTIC (APPROXIMATE).
GLUCAN-BINDING (APPROXIMATE).
A REPEAT.
5 X TANDEM REPEATS.
1.
2.
3.
4.
5.
S -> T (IN STRAIN MT4239).
T -> I (IN STRAIN GS-5).
V -> A (IN STRAINS GS-5, MT4245, MT4251,
MT4467 AND MT8148).
O -> P (IN STRAIN MT4251).
I -> S (IN STRAINS GS-5, MT4245, MT4251,
MT4467 AND MT8148).
S -> F (IN STRAIN MT4251).
K -> N (IN STRAIN MT4251).
S -> D (IN STRAINS GS-5, MT4467 AND
MT8148).
N -> R (IN STRAIN MT4239).
I -> T (IN STRAIN MT4239).
K -> R (IN STRAIN MT8148).
F -> Y (IN STRAIN MT8148).
T -> I (IN STRAIN MT8148).
A -> V (IN STRAIN MT8148).
F -> L (IN STRAIN MT8148).
FGEFVE -> YGTPVA (IN STRAINS GS-5, MT4239
AND MT4467).
SV -> NT (IN STRAINS GS-5, MT4239 AND
MT4467).
ADS -> VDG (IN STRAINS GS-5, MT4239 AND
MT4467).
A -> T (IN STRAIN MT4239).
S -> N (IN STRAIN MT4239).
H -> Y (IN STRAIN MT4251).
E -> K (IN STRAIN MT8148).
Y -> C (IN STRAIN MT8148).
A -> P (IN STRAIN MT4239).
R -> P (IN STRAIN MT8148).
R -> H (IN STRAINS GS-5 AND MT4467).
Y -> H (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
S -> G (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
H -> Y (IN STRAINS GS-5 AND MT4467).
Y -> H (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
S -> G (IN STRAINS GS-5, MT4239, MT4467

FT AND MT8148).
FT VARIANT 1459 1459 Y -> H (IN STRAIN MT4467).
FT CONFLICT 570 570 R -> A (IN REF. 1).
FT CONFLICT 800 817 ADDVAVAASTASTDGG -> LKKFAURLARPHQQA
(IN REF. 1).
FT CONFLICT 1310 1310 H -> L (IN REF. 1).
SQ SEQUENCE 1476 AA; 165846 MW; 9C6E09F731B4CBCE CRC64;
Query Match 100.0%; Score 119; DB 1; Length 1476;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANDHSLTEAWSDNDTPYLHD 21
Db 481 ANDHSLTEAWSDNDTPYLHD 501
RESULT 2
ID GTFC STRNU STANDARD; PRT; 1455 AA.
AC P13470; O69382; O69385; O69388; O69391; O69397; P05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-ST precursor (EC 2.4.1.5) (GTF-ST)
GN (dextranucrase) (sucrose 6-glucosyltransferase).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfc gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
MT4467 / Serotype E and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9770124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Ferretti J.J.;
RA Li S., Zhu H., Najjar F., Lai H., White T., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
[4]
RP SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfc gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.

```

FT FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148) .
FT FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239) .
FT FT VARIANT 1439 1439 V -> I (IN STRAINS MT1429 AND MT8148) .
FT FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148) .
FT FT CONFLICT 1337 1455 ORLYFSGNGVQAKCELITTEKRGRTKYIDPNSGNREVRNRYR
TT TSSGNNWYFGNDGYALIGMHVBEGRVYFDENGVRTASHD
PT ORNMHDYDRRDTCRGSSSAVRFHRSNGFPDNFRF ->
FT HASILSMVERLRRESSLQSIVSVNTMILLPEMKFVIWM
FT (IN REF. 1) .
FT FT SEQUENCE 1455 AA; 162965 MW; 3C455A994FEC86 CRC64;
SQ QUERY MATCH 92.4%; Score 110; DB 1; Length 1455;
Best Local Similarity 95.2%; Pred. No. 3,38-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
QY 1 ANDHLSLEAWSDNDPYLHD 21
Db 507 ANDHLSLEAWSYNDPYLHD 527
RESULT 3
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE GLUCOSYLTRANSFERASE-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase) .
OS Streptococcus downei (Streptococcus sobrinus) .
OC Streptia; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1317;
[1]
SEQUENCE FROM N.A.
RP STRAIN=6715 / Serotype G;
RC MEDLINE=91123227; PubMed=1704006;
RX Abo H., Matsumura T., Kodama T., Ohca H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase)";
RL J. Bacteriol. 173:989-996(1991).
CC -I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDiate THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -I- CATALYTIC ACTIVITY: Sucrose + ({1,6}-alpha-D-glucosyl) (N) = D-
CC fructose + ({1,6}-alpha-D-glucosyl) (N+1).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DISEASE: DENTAL CARIES.
CC -I- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -I- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D90213; BAAL4241.1; --
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro..70.
DR Pfam; PF01473; CW_binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.

```

```

FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-1.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351BCF CRC64;

Query Match 90.8%; Score 108; DB 1; Length 1592;
Best Local Similarity 85.7%; Pred. No. 7.6e-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAMSNDTPYLHD 21
DB 477 ANNHVSIVEAMSNDTPYLHD 497

RESULT 4
GTF1_STRDO STANDARD; PRT: 1597 AA.
ID GTF1_STRDO
AC P11001.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFE28."
RL J. Bacteriol. 169:4271-4278(1987).
CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC - CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
CC fructose + (1,6)-alpha-D-glucosyl (N+1).
CC - SUBCELLULAR LOCATION: Secreted.
CC - DISEASE: DENTAL CARIES.
CC - MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC - SIMILARITY: Contains 19 cell wall binding repeats.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17391; AAC63063.1;
CC InterPro; IPR002479; CW_binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CW_binding_1; 16.

```

```

DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-1.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1597 1.
FT REPEAT 1163 1213 A REPEAT.
FT REPEAT 1227 1273 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 90.8%; Score 108; DB 1; Length 1597;
Best Local Similarity 85.7%; Pred. No. 7.6e-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAMSNDTPYLHD 21
DB 483 ANNHVSIVEAMSNDTPYLHD 503

RESULT 5
GTF5_STRDO STANDARD; PRT: 1365 AA.
ID GTF5_STRDO
AC P29336.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTF5.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell R.R., Ferretti J.J.;
RT "Analysis of the Streptococcus downei gtfS gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans."
RL Infect. Immun. 58:2452-2458(1990).
CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC - CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
CC fructose + (1,6)-alpha-D-glucosyl (N+1).
CC - ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC - DISEASE: DENTAL CARIES.
CC - MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC - SIMILARITY: Contains 10 cell wall binding repeats.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M30943; AAA26898.1;

```

DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 36
 FT CHAIN 37 1365
 FT REPEAT 157 177
 FT REPEAT 178 197
 FT REPEAT 198 1061
 FT REPEAT 1062 1082
 FT REPEAT 1083 1102
 FT REPEAT 1150 1169
 FT REPEAT 1170 1190
 FT REPEAT 1225 1243
 FT REPEAT 1289 1308
 FT REPEAT 1309 1328
 FT REPEAT 1331 1352
 SQ SEQUENCE 1365 AA; 151590 MW; 16729685A2E8C476 CRC64;

Query Match 65.5%; Score 78; DB 1; Length 1365;
 Best Local Similarity 71.4%; Pred. NO. 0.00037;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAMSDNDPYLHD 21
 Db 467 AIDHLSTLEAMSGNDNDYVD 487

RESULT 6
 GTFD_STRMU STANDARD; PRT: 1462 AA.
 AC P49331; O69383; O69386; O69389; O69392; O69398;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFD OR SMU.910.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId=1309;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=91100958; PubMed=2148600;
 RA Honda O., Kato C., Kuramitsu H.K.;
 RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
 the glucosyltransferase-S enzyme";
 RL J. Gen. Microbiol. 136:2099-2105(1990).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 RC MT4467 / Serotype E, and MT8148 / Serotype C;
 RX MEDLINE=98231643; PubMed=9570124;
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of
 Streptococcus mutans";
 RL FEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Iln S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -!- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
 CC fructose + {(1,6)-alpha-D-glucosyl}(N+1).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DISEASE: DENTAL CARIES.
 CC -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S.MUTANS.
 CC -!- SIMILARITY: Contains 6 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M29286; AAA26895.1; -
 CC EMBL: D86653; BAA26103.1; -
 CC EMBL: D86656; BAA26107.1; -
 CC EMBL: D86659; BAA26111.1; -
 CC EMBL: D86662; BAA26115.1; -
 CC EMBL: D86979; BAA26121.1; -
 CC EMBL: A6014932; AAN58619.1; -
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 FT SIGNAL 1 1462
 FT CHAIN 1232 1423
 FT DOMAIN 1232 1295
 FT REPEAT 1296 1359
 FT REPEAT 1360 1423
 FT REPEAT 10 10
 FT VARIANT 19 19
 FT VARIANT 58 58
 FT VARIANT 68 68
 FT VARIANT 81 81
 FT VARIANT 113 113
 FT VARIANT 122 122
 FT VARIANT 132 132
 FT VARIANT 135 135
 FT VARIANT 137 137
 FT VARIANT 202 202
 FT VARIANT 255 255
 FT VARIANT 275 275
 FT VARIANT 288 288
 FT VARIANT 301 301
 FT VARIANT 313 313
 FT VARIANT 317 317
 FT VARIANT 328 328
 FT VARIANT 350 350
 FT VARIANT 628 633
 FT VARIANT 688 688
 FT VARIANT 726 732
 FT VARIANT 726 730
 FT VARIANT 762 762

POTENTIAL.
 GLUCOSYLTRANSFERASE-S.
 3 X 63 AA APPROXIMATE TANDEM REPEATS.
 1.
 2.
 3.
 Y -> H (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148)
 I -> V (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 K -> E (IN STRAIN MT4467).
 A -> S (IN STRAINS MT4239 AND MT4245).
 A -> T (IN STRAINS MT4251 AND MT8148).
 I -> V (IN STRAINS MT4239, MT4245 AND
 MT8148).
 A -> S (IN STRAINS GS-5 AND MT4467).
 A -> V (IN STRAIN MT4245).
 A -> T (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 V -> L (IN STRAIN MT4239).
 D -> N (IN STRAIN MT8148).
 E -> D (IN STRAINS MT4239, MT4245 AND
 MT4251).
 D -> N (IN STRAINS MT4239, MT4245 AND
 MT4251).
 Q -> H (IN STRAIN MT4245).
 D -> N (IN STRAINS GS-5 AND MT4251).
 E -> K (IN STRAIN MT4239).
 V -> F (IN STRAIN MT4239).
 F -> L (IN STRAINS MT4239, MT4251 AND
 MT4467).
 KKKTQ -> EKETL (IN STRAIN MT4251).
 A -> S (IN STRAIN MT4239).
 TDQSEA -> ADKGNS (IN STRAIN MT4251).
 TDQGS -> ADKGN (IN STRAINS MT4239 AND
 MT4245).
 T -> A (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).

```

FT VARIANT 964 964 D -> Y (IN STRAIN MT4251).
FT VARIANT 1019 1019 E -> K (IN STRAINS MT4245 AND MT4251).
FT VARIANT 1059 1060 LG -> R (IN STRAIN MT4251).
FT VARIANT 1060 1060 G -> R (IN STRAIN MT4245).
FT VARIANT 1080 1080 G -> R (IN STRAIN MT4245).
FT VARIANT 1142 1142 H -> Q (IN STRAIN GS-5).
FT VARIANT 1142 1142 S -> N (IN STRAIN MT4239).
FT VARIANT 1198 1198 Y -> C (IN STRAINS MT4251 AND MT4467).
FT VARIANT 1220 1220 F -> L (IN STRAIN MT4467).
FT VARIANT 1280 1280 Q -> P (IN STRAIN MT4245).
FT VARIANT 1282 1282 K -> T (IN STRAIN MT4245).
FT VARIANT 1290 1290 N -> D (IN STRAIN MT4245).
FT VARIANT 1311 1311 G -> D (IN STRAINS GS-5 AND MT4467).
FT VARIANT 1403 1403 G -> R (IN STRAIN GS-5).
FT VARIANT 1425 1425 R -> K (IN STRAIN MT4467).
FT VARIANT 1449 1449 RYDKNSGMVYKWTIANGRIIGDIRMGIRARY -> VY
FT CONFLICT 1428 1462 R (IN REF. 1).
SQ SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;

Query Match 63.0%; Score 75; DB 1; Length 1462;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 1 ANDHSTIEAMSDNDTPYLH 21
Db 495 AINHLSTIEAMSDNDPOYNKD 515

RESULT 7
G3PC_LEIME STANDARD; PRT; 330 AA.
ID G3PC_LEIME
AC Q01558;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (GAPDH).
GN GAPC.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_Taxid=5665;
RA Hamaert V., Blaauw M., Kohl L., Allert S., Opperdoes F.R.,
RA Michels P.A.M.;
RT "Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-
RT phosphate dehydrogenase in Leishmania mexicana.";
RT Mol. Biochem. Parasitol. 55:115-126(1992).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PARTWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X65220; CAA46323.1; --
CC PIR; B48445; B48445.
CC HSSP; P06977; IGAD.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh_1.
CC Pfam; PF02800; gpdh_C_1.

```

```

DR PRINTS; PR00078; G3PDHDEGNASE.
DR TIGRfams; TIGR01534; GAPDH-I. 1.
DR PROSITE; PS00071; GAPDH; 1.
KM Glycolysis; Oxidoreductase; NAD.
FT INIT MET 0 BY SIMILARITY.
FT BINDING 148 148 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 175 175 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 330 AA; 35511 MW; ED46BD88A07F1E CRC64;

Query Match 47.9%; Score 57; DB 1; Length 330;
Best Local Similarity 52.6%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Oy 2 NDHSTIEAMSDNDTPYLH 20
Db 300 NDHFWKLVSVYNDHETGYSH 318

RESULT 8
YBHK_SALTI STANDARD; PRT; 302 AA.
ID YBHK_SALTI
AC P58587;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
GN YBHK OR STY0835 OR T2086.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=601;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leach R.S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Skellerns M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RT Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=TY2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burdand V., Kodoyianni V., Schwartz D.C., Blatner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RT J. Bacteriol. 185:2330-2337(2003).
CC -1- SIMILARITY: Belongs to the UPF0052 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL627268; CAD05249.1; --
CC DR EMBL; AE016841; AAO69704.1; --
CC DR InterPro; IPR02882; UPF0052.
CC DR Pfam; PF01933; UPF0052; 1.
CC KM Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 302 AA; 32656 MW; 2218471211E185C3 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 302;

```

Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDPYLEHD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 9

YBHK_SALTY STANDARD; PRT; 302 AA.

AC P58586;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
GN YBHK OR STM0801.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxId=602;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=1677609;
RA McClelland M., Sanderson K.R., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulyanay E.,
RA Ryan E., Sun H., Flores R., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC - SIMILARITY: Belongs to the UPF0052 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AB008733; AAL19738.1;
DR StyGene; SG32227; ybhK.
DR InterPro: IPR002882; UPF0052.
DR Pfam; PF01933; UPF0052; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32626 MW; 38A84708A1B185C3 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 302;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDPYLEHD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 10
UL16_EBV STANDARD; PRT; 336 AA.

AC P303231;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE BGJF2 protein.
GN BGJF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxId=10377;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seglin C.,
RA Tuffnell P.S., Watnall B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113548; PubMed=1662696;
RA Chen M.R., Hsu T.Y., Lin S.W., Chen J.Y., Yang C.S.;
RT "Cloning and characterization of cDNA clones corresponding to
RT transcripts from the BamHI G region of the Epstein-Barr virus genome
RT and expression of BGJF2.";
RL J. Gen. Virol. 72:3047-3055(1991).
CC - SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
CC HSV-6 ORF11R, EBV-1 46, HCMV UL94, EBV BDLF2, AND VZV 44.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M60514; AAA45871.1;
DR EMBL: V01555; CAA24831.1;
DR EMBL: S77132; AAB21113.1;
DR PIR; C43044; Q0B840.
DR InterPro: IPR004286; UL16_UL94.
DR Pfam; PF03044; UL16_UL94; 1.
KM Late protein.
SQ SEQUENCE 336 AA; 36888 MW; 840937A416D5584C CRC64;

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 2.8;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDP 17
DB 216 AGAHVILKGMITDSDP 232

RESULT 11
G3PC_TRYB STANDARD; PRT; 330 AA.
AC P10097;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (GAPDH).

OS Trypanosoma brucei brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI_TaxId=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=427;
RX MEDLINE=91249838; PubMed=2040303;
RA Michels P.A.M., Marchand M., Kohl L., Allert S., Wierenga R.K.,
RT "The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate
RT dehydrogenase in Trypanosoma brucei have a distant evolutionary
RT relationship.";
RL Eur. J. Biochem. 198:421-428(1991).

RN [2]
RP PRELIMINARY SEQUENCE OF 1-85.
RX MEDLINE=87161817; PubMed=3830153;
RA Misset O., van Beeumen J., Lambel A.M., van der Meer R.,
RA Opperdoes F.R.;
RT "Glyceraldehyde-phosphate dehydrogenase from Trypanosoma brucei.
RT Comparison of the glycosomal and cytosolic isoenzymes.";


```

RL Eur. J. Biochem. 162:501-507(1997).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis, first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53472; CAA37568.1; -.
CC PIR: S16091; DEUTGC.
CC HSSP: P06977; 1GMD.
CC InterPro: IPR000173; GAP_dhhydrogenase.
CC DR InterPro: IPR006424; GAPDH-I.
CC DR Pfam: PF00044; gpdh. 1.
CC DR Pfam: PF02800; gpdh_C. 1.
CC DR PRINTS: PR00078; G3PDHGRGNASE.
CC DR TIGRFS: TIGR01534; GAPDH-I; 1.
CC DR PROSITE: PS00071; GAPDH. 1.
CC KW Glycolysis; Oxidoreductase; NAD.
CC INIT_MET 0
CC BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
CC ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC VARIANT 72 72 R -> I.
CC CONFLICT 56 60 Q -> T (IN REF. 2).
CC CONFLICT 60 60 Q -> K (IN REF. 2).
CC FT SEQUENCE 330 AA; 35503 MW; 7E1F0AFD4091E7C5 CRC64;
SQ
Query Match 39.9%; Score 47.5; DB 1; Length 330;
Best Local Similarity 43.5%; Pred. No. 4.7;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;
QY 2 NDHSLILEAWSNDTPY---LHD 21
Db 301 NDNFVKLVSWYDNETGYSNKVD 323
RESULT 12
G3P_COCPO STANDARD; PRT; 337 AA.
AC Q8J1H3;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GPD.
OS Coccidioides posadasii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Omygenales; mitosporic Omygenales; Coccidioides.
OX NCBI_TaxID=199306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C735;
RA Hung C.-Y., Yu J.-J., Cole G.T.;
RT "Isolation and evaluation of expression of the glyceraldehyde-3-
RT phosphate dehydrogenase gene of Coccidioides immitis."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis, first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF288134; AAN76496.1; -.
CC InterPro: IPR006424; GAP_dhhydrogenase.
CC DR InterPro: IPR000173; GAP_dhhydrogenase.
CC DR Pfam: PF00044; gpdh. 1.
CC DR Pfam: PF02800; gpdh_C. 1.
CC DR PRINTS: PR00078; G3PDHGRGNASE.
CC DR TIGRFS: TIGR01534; GAPDH-I; 1.
CC DR PROSITE: PS00071; GAPDH. 1.
CC KW Glycolysis; Oxidoreductase; NAD.
CC BINDING 151 151
CC ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC FT SEQUENCE 337 AA; 36400 MW; 62EA834CA536DAF2 CRC64;
SQ
Query Match 39.5%; Score 47; DB 1; Length 337;
Best Local Similarity 52.9%; Pred. No. 5.7;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 NDHSLILEAWSNDTPY 18
Db 303 NDNFVKLVSWYDNETGYSNKVD 319

```

```

RESULT 13
SC35_YEAST STANDARD; PRT; 275 AA.
AC P53271;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SEC35 protein.
GN SEC35 OR YGR120C OR G6324.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RC MEDLINE=97197982; PubMed=9046098;
RA van Dyck L., Tetteijn H., Purnelle B., Goffeau A.;
RT "An 18.3 kb DNA fragment from yeast chromosome VII carries four
RT unknown open reading frames, the gene for an Aen synthase, remnants
RT of Ty and three tRNA genes."
RL Yeast 13:171-176(1997).
RN [2]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RN MEDLINE=98270888; PubMed=9606204;
RA Vanheenen S.M., Cao X., Japashin V.V., Barlowe C., Waters M.G.;
RT "Sec35p, a novel peripheral membrane protein, is required for ER to
RT Golgi vesicle docking."
RL U. Cell Biol. 141:1107-1119(1998).
RN [3]
RP SUBUNIT.
RN MEDLINE=21563418; PubMed=11703943;
RA Whyte J.R., Munro S.;
RT "The Sec34/35 Golgi transport complex is related to the exocyst,
RT defining a family of complexes involved in multiple steps of membrane
RT traffic."
RL Dev. Cell 1:527-537(2001).
CC -1- FUNCTION: Required for ER to Golgi vesicle docking. Not essential
CC for viability.
CC -1- SUBUNIT: Component of the Sec34/Sec35 complex which consists of
CC eight different proteins.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X83099; CAA58155.1; -
 DR EMBL; 272905; CAA97130.1; -
 DR PIR; S55978; S55978.
 DR SGD; S0003352; SEC35.
 DR GO; GO:0017119; C:Golgi transport complex; IPI.
 DR GO; GO:0000300; C:peripheral membrane protein of membrane fra. .; IDA.
 DR GO; GO:0005625; C:soluble fraction; IDA.
 DR GO; GO:0005515; F:protein binding activity; IPI.
 DR GO; GO:0006888; P:ER to Golgi transport; IGI.
 DR GO; GO:0000301; P:retrograde vesicle recycling within Golgi. .; IMP.
 KW Transport; Protein transport; Golgi stack; Membrane.
 SQ SEQUENCE 275 AA; 31799 MW; AA102D086FF37AD7 CRC64;

Query Match
 Best Local Similarity 38.7%; Score 46; DB 1; Length 275;
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 7 ILEAMSDNTPYLH 20
 DB 217 ILESCANSNPYIH 230

RESULT 14
 G3P_CHLPPN STANDARD; PRT; 335 AA.
 ID G3P_CHLPPN
 AC Q92770; Q9Q0H7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR GAPA OR CPN0624 OR CP0123.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 NCBI_TaxID=83558;
 RX MEDLINE=9206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Titterack T., Berry K., Bass S.,
 RA Linher K., Weidman J., Kouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBUNITAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE001647; AAD18763.1; -
 DR EMBL; AE002173; AAF38006.1; -
 DR EMBL; AP002547; BAA39831.1; -
 DR PIR; B72053; B72053.
 DR PIR; B86568; B86568.
 DR HSSP; P06977; IGAD.
 DR PHCI-2DPAGE; Q92770; -
 DR TIGR; CP0123; -
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; spdh.1.
 DR Pfam; PF02800; spdh.C.1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH.1.
 KW Glycolysis; Oxidoreductase; NAD; Complete proteome
 FT BINDING 151
 FT 151
 FT ACT_SITE 178 178
 FT ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).

SQ SEQUENCE 335 AA; 36837 MW; C86D0B3ADAD8FE CRC64;

Query Match
 Best Local Similarity 38.7%; Score 46; DB 1; Length 335;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNTPY 18
 DB 303 NDRFFELVAMVDNENRGY 319

RESULT 15
 PCHA_PSEAE STANDARD; PRT; 476 AA.
 ID PCHA_PSEAE
 AC Q51508;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salicylate biosynthesis isochorismate synthase (EC 5.4.99.6)
 DE (isochorismate mutase).
 GN PCHA OR PA4231.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=96086939; PubMed=7500944;
 RA Serino L., Reimann C., Baur H., Beyeler M., Visca P., Haas D.;
 RT "Structural genes for salicylate biosynthesis from chorismate in
 RT Pseudomonas aeruginosa.";
 RL Mol. Gen. Genet. 249:217-228(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lozy S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen." ;
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF CHORISMATE TO SALICYLATE
 CC (PROBABLE).
 CC -1- CATALYTIC ACTIVITY: Chorismate = isochorismate.
 CC -1- PATHWAY: Salicylic acid biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ISOCHORISMATE SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X82644; CAAS7969.1; -;
 DR EMBL; AE004840; AAG07619.1; -;
 DR PIR; S60203; S58229.
 DR InterPro: IPR005801; Anth synth chor.
 DR InterPro: IPR004561; Isochor syn.
 DR Pfam; PF00425; chorismate_bind; 1.
 DR ProDom; PD000779; Anth_synth_chor; 1.
 DR TrEMBL; TIGR00543; isochor_syn; 1.
 KW isomerase; Complete proteome.
 KW SEQUENCE 476 AA; 52071 MW; D0F6AFD9CDCF9CF5 CRC64;
 SQ
 Query Match 38.7%; Score 46; DB 1; Length 476;
 Best local Similarity 44.4%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 Oy 3 DHTSILEAFMSDNDTPYLH 20
 Db 44 DPLOVFGAWDDDRQTPTCY 61

Search completed: November 13, 2003, 09:45:26
 Job time : 8.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-1

Perfect score: 119

Sequence: 1 ANDHLSTLEAWSDNDTPYLHD 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR 76:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	119	100.0	1475	2 B33135	glfB protein precu
2	110	92.4	1375	2 J70345	dextranucrase (EC
3	108	90.8	1592	2 A38175	glucosyltransferas
4	78	65.5	1365	2 A41483	glucosyltransferas
5	75	63.0	1431	2 A45866	dextranucrase (EC
6	74	62.2	1508	2 T31098	probable dextranu
7	66	55.5	1449	2 T30857	glucosyltransferas
8	63	52.9	1577	2 T30858	glucosyltransferas
9	58	48.7	1518	2 A44811	glucosyltransferas
10	58	47.9	1331	2 B48445	glycerinaldehyde-3-p
11	57	47.9	1599	2 S22737	glucosyltransferas
12	57	47.9	1599	2 T30289	pristinamycin 1 sy
13	52	43.7	4848	2 T48610	hypothetical prote
14	50	42.0	347	2 T03983	rif2 nuclear resor
15	49.5	41.6	549	2 AG0597	conserved hypotet
16	49	41.2	302	2 AG0597	conserved hypotet
17	49	41.2	336	2 Q08A40	BGLF2 protein - hu
18	49	41.2	601	2 B87028	pyruvate, phosphat
19	48	40.3	490	2 H70538	probable pdk prot
20	47.5	39.9	331	1 DEUTGC	glycerinaldehyde-3-p
21	47.5	39.9	519	2 C86372	hypothetical prote
22	47.5	39.9	175	2 C86205	hypothetical prote
23	47	39.5	378	2 T04254	hypothetical prote
24	47	39.5	524	2 D82220	conserved hypotet
25	47	39.5	525	2 T40088	RNaseG domain cont
26	46.5	39.1	418	2 D90506	4-aminobutylate am
27	46	38.7	275	2 S55978	hypothetical prote
28	46	38.7	335	2 B86568	glycerinaldehyde-3-p
29	46	38.7	335	2 B72053	glycerinaldehyde-3-p

30	46	38.7	400	2 F98138	hypothetical prote
31	46	38.7	417	2 AE3149	glucosyltransferas
32	46	38.7	476	1 S58229	salicylate biosynt
33	45.5	38.2	96	2 E81786	conserved hypotet
34	45	37.8	148	2 B95908	guanydinocetate N
35	45	37.8	236	2 S62732	conserved hypotet
36	45	37.8	309	2 D95905	conserved hypotet
37	45	37.8	947	2 E86362	hypothetical prote
38	45	37.8	1090	2 AG1745	glycosidase homolo
39	45	37.8	1091	2 AF1380	glycosidase homolo
40	45	37.8	1159	2 T43461	probable phosphodi
41	44.5	37.4	210	2 H83332	conserved hypotet
42	44	37.0	155	2 AD2392	transcription regu
43	44	37.0	247	2 PC0178	glycerinaldehyde-3-p
44	44	37.0	331	1 DEECG3	glycerinaldehyde-3-p
45	44	37.0	331	1 H90939	glycerinaldehyde-3-p

ALIGNMENTS

RESULT 1
B33135
glfB protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C:Accession: B33135; A33128
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the glfB gene from Streptococcus mutans.
A:Reference number: A33135; PMID:87308013; PMID:3040685
A:Accession: B33135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1475 <SH1>
A:Cross-references: GB:M17361; NID:G153639; PIDN:AA88588.1; PID:G153640
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990
A:Reference number: A33128
A:Accession: A33128
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-171,173-641,'N',643-1475 <SH2>
A:Experimental source: strain GS-5
A:Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 100.0%; Score 119; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5,2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSTLEAWSDNDTPYLHD 21
DB 481 ANDHLSTLEAWSDNDTPYLHD 501

RESULT 2
J70345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N:Alternate names: Sucrose 6-glucosyltransferase
C:Species: Streptococcus mutans
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C:Accession: J70345; C33135
R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988.
A:Title: Sequence analysis of the glfC gene from Streptococcus mutans GS-5.
A:Reference number: J70345; PMID:89137980; PMID:2976010
A:Accession: J70345
A:Molecule type: DNA

A:Residues: 1-1375 <UED>
A:Experimental source: GS-5
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the *gtfB* gene from *Streptococcus mutans*.
A:Reference number: A33155; MUID:87308013; PMID:3040685
A:Accession: G33135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <SHI>
A:Cross-references: GB:M17361
C:Genetics:
A:Gene: *gtfC*
C:Function:
A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C:Superfamily: cpl repeat homology
C:Keywords: duplication; glycosyltransferase; hexosyltransferase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-1375/Product: glycosyltransferase #status predicted <MAT>
F:1126-1145/Domain: cpl repeat homology <CP1>
F:1253-1272/Domain: cpl repeat homology <CP2>
F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 92.4%; Score 110; DB 2; Length 1375;
Best Local Similarity 95.2%; Pred. No. 1,2e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSGNDPTPLMD 21
DB 507 ANDHSLILEAMSGNDPTPLMD 527

RESULT 3

A38175
glucosyltransferase precursor - *Streptococcus sobrinus*
C:Species: *Streptococcus sobrinus*
C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1995
C:Accession: A38175
R:Abdo, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
J. Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within *Streptococcus*
A:Reference number: A38175; MUID:91123227; PMID:1704006
A:Accession: A38175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1592 <ABO>
A:Cross-references: GB:D90213; NID:9217032; PIDN:BA14241.1; PID:dl014946; PID:9217033
C:Superfamily: cpl repeat homology
F:1093-1112/Domain: cpl repeat homology <CP1>
F:1222-1241/Domain: cpl repeat homology <CP2>
F:1287-1306/Domain: cpl repeat homology <CP3>
F:1330-1351/Domain: cpl repeat homology <CP4>
F:1352-1371/Domain: cpl repeat homology <CP5>
F:1402-1420/Domain: cpl repeat homology <CP6>
F:1465-1484/Domain: cpl repeat homology <CP7>
F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 90.8%; Score 108; DB 2; Length 1592;
Best Local Similarity 85.7%; Pred. No. 3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSGNDPTPLMD 21
DB 477 ANDHSLILEAMSGNDPTPLMD 497

RESULT 4

A41483
glucosyltransferase (EC 2.4.1.-) *gtfs* precursor - *Streptococcus sobrinus*
C:Species: *Streptococcus sobrinus*
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1995
C:Accession: A41483
R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

Infect. Immun. 58, 2452-2458, 1990
A:Title: Analysis of the *Streptococcus downei* *gtfs* gene, which specifies a glucosyltransferase
A:Reference number: A41483; MUID:90316665; PMID:2142479
A:Accession: A41483
A:Molecule type: DNA
A:Residues: 1-1365 <GIL>
A:Cross-references: GB:M30943; NID:9153652; PIDN:AAA26898.1; PID:9153653
C:Genetics:
A:Gene: *gtfs*
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
Query Match 65.5%; Score 78; DB 2; Length 1365;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSGNDPTPLMD 21
DB 467 ANDHSLILEAMSGNDPTPLMD 487

RESULT 5

A45866
dextranucrase (EC 2.4.1.5) precursor - *Streptococcus mutans*
C:Species: *Streptococcus mutans*
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A45866
R:Honda, O.; Kato, C.; Kuramitsu, H. K.
J. Gen. Microbiol. 136, 2099-2105, 1990
A:Title: Nucleotide sequence of the *Streptococcus mutans* *gtfd* gene encoding the glucosyltransferase
A:Reference number: A45866; MUID:91100958; PMID:2148600
A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1129-1211/Domain: cpl repeat homology <CP3>
F:1257-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP7>
F:1385-1404/Domain: cpl repeat homology <CP8>

Query Match 63.0%; Score 75; DB 2; Length 1431;
Best Local Similarity 71.4%; Pred. No. 0.0036;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSGNDPTPLMD 21
DB 495 ANDHSLILEAMSGNDPTPLMD 515

RESULT 6

T31098
probable dextranucrase (EC 2.4.1.5), extracellular - *Leuconostoc mesenteroides*
C:Species: *Leuconostoc mesenteroides*
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T31098
R:Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R. M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A:Reference number: Z20981; MUID:98164374; PMID:9503626
A:Accession: T31098
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1508 <MON>
A:Cross-references: EMBL:AF030129; NID:92766611; PID:92766612; PIDN:AA95453.1
A:Experimental source: strain NRRL B-1299
C:Genetics:

A/Gene: dsrB
A/Function: produces dextran composed only of alpha(1-6) glucosidic bonds
A/Description: Streptococcus salivarius
C/Keywords: glycosyltransferase, hexosyltransferase

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.0055;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLIEMSDNDTPYLHD 21
DB 563 ANOHLIEMSDNDPEYKD 583

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius
C/Species: Streptococcus salivarius
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Accession: T30857

R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A/Reference number: Z20909; MUID:95122197; PMID:7822030
A/Accession: T30857
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-1449 <SIM>
A/Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AA041412.1
C/Keywords: glycosyltransferase, hexosyltransferase
A/Gene: gtfB

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLIEMSDNDTPYLHD 21
DB 539 HSLIEMSDNDAYNED 556

RESULT 8
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)

C/Species: Streptococcus salivarius
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Accession: T30552
R/Jaffe, R.I.
submitted to the EMBL Data Library, February 1998

A/Description: Streptococcus salivarius V1477 gtfN.
A/Reference number: Z20854
A/Accession: T30552
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-1449 <JAF>
A/Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AA05155.1
C/Keywords: glycosyltransferase, hexosyltransferase
A/Gene: gtfN

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLIEMSDNDTPYLHD 21
DB 539 HSLIEMSDNDAYNED 556

RESULT 9
T30858
glucosyltransferase - Streptococcus salivarius
C/Species: Streptococcus salivarius
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C/Accession: T30858
R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995
A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A/Reference number: Z20909; MUID:95122197; PMID:7822030
A/Accession: T30858
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1577 <SIM>
A/Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIDN:AA041413.1
C/Keywords: glycosyltransferase, hexosyltransferase
A/Gene: gtfB

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.3;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLIEMSDNDTPYLHD 21
DB 591 HSLIEMSDNDHQNKD 608

RESULT 10
A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius

C/Species: Streptococcus salivarius
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C/Accession: A44811; S22726; S28809
R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991
A/Title: Molecular characterization of a cluster of at least two glucosyltransferase genes

A/Reference number: A44811; MUID:92148377; PMID:1838391
A/Accession: A44811
A/Molecule type: DNA
A/Residues: 1-1518 <GIF>

A/Cross-references: EMBL:U11873; NID:g47526; PIDN:CAA77900.1; PID:g47527
A/Note: Sequence extracted from NCBI backbone (NCBIN:81050, NCBI:81052)
C/Keywords: glycosyltransferase, hexosyltransferase
A/Gene: gtfB

C/Superfamily: cpl repeat homology
C/Keywords: glycosyltransferase, hexosyltransferase
F,1307-1326/Domain: cpl repeat homology <CP4>

Query Match 48.7%; Score 58; DB 2; Length 1518;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLIEMSDNDTPY 18
DB 504 HSLIEMSDNDNH 518

RESULT 11

B48445
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania mex

C/Species: Leishmania mexicana
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C/Accession: B48445; S25142
R/Haeneert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Opperdoes, F.R.; Michels, P.A.M.

Mol. Biochem. Parasitol. 55, 115-126, 1992
A/Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate de

A/Reference number: A48445; MUID:93063042; PMID:1435864
A/Accession: B48445
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-331 <HAN>
A/Cross-references: EMBL:X65220; NID:g9552; PIDN:CAA46323.1; PID:g9553
C/Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C/Keywords: oxidoreductase

Query Match 47.9%; Score 57; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NDHSLLEAWSNDPTPLH 20
DB 301 NDHFVKLWSWIDNETGYSH 319

RESULT 12

Glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
C/Species: Streptococcus salivarius
C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C/Accession: S22737; S28810; B44811; S22727
R/Jacques, N.
submitted to the EMBL Data Library, March 1992
A/Reference number: S22726
A/Accession: S22737
A/Molecule type: DNA
A/Residues: 1-1599 <UAC>
A/Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAA77898.1; PID:G47531
A/Experimental source: ATCC 25975
R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A/Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
A/Reference number: A44811; MUID:92148377; PMID:1838391
A/Accession: S28810
A/Molecule type: DNA
A/Residues: 1-51 <GIF>
A/Cross-references: EMBL:Z11873
C/Genetics:
A/Status: preliminary
C/Suprafamily: cpl repeat homology
C/Keywords: glycosyltransferase; hexosyltransferase
F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 47.9%; Score 57; DB 2; Length 1599;
Best Local Similarity 55.6%; Pred. No. 2.6;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLLEAWSNDPTPLH 21
DB 494 NLSLEAWSNDPTPLH 511

RESULT 13
T30289
pristinamycin I synthase 3 - Streptomyces pristinaespiralis
C/Species: Streptomyces pristinaespiralis
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Dec-2000
C/Accession: T30289
R/Creevy-Lagard, V.A.; Saurin, W.; Thibaut, D.; Gil, P.; Naudin, L.; Crouzet, J.; Bian
submitted to the EMBL Data Library, February 1997
A/Description: Streptogramin B biosynthesis in Streptomyces pristinaespiralis and Strept
A/Reference number: Z20808
A/Accession: T30289
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4848 <DEC>
A/Cross-references: EMBL:Y11548; NID:e1025755; PID:e307539; PIDN:CAA72312.1
C/Genetics:
A/Status: preliminary
C/Suprafamily: acyl carrier protein homology; acetate-CoA ligase homology
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:511-951/Domain: acetate-CoA ligase homology <ACLI>
F:966-1034/Domain: acyl carrier protein homology <ACPI>
F:1563-2024/Domain: acetate-CoA ligase homology #status atypical <ACLI2>
F:2399-2467/Domain: acyl carrier protein homology <ACP2>
F:2995-3427/Domain: acetate-CoA ligase homology <ACL3>
F:3441-3509/Domain: acyl carrier protein homology <ACP3>
F:4043-4492/Domain: acetate-CoA ligase homology <ACL4>
F:4507-4575/Domain: acyl carrier protein homology <ACP4>
F:998,2431,3473/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 43.7%; Score 52; DB 2; Length 4848;

Best Local Similarity 47.1%; Pred. No. 58;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLLEAWSNDPTP 17
DB 3972 ADHARVLGWNDDHP 3988

RESULT 14

hypothetical protein F18022.160 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T48610
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24493
A/Accession: T48610
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-347 <BEV>
A/Cross-references: EMBL:DB
A/Experimental source: cultivar Columbia; BAC clone F18022
C/Genetics:
A/Map position: 5
A/Intons: 232/3; 296/3
A/Note: F18022.160

Query Match 42.0%; Score 50; DB 2; Length 347;
Best Local Similarity 52.9%; Pred. No. 5.2;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 DHSLEAWSNDPTPL 19
DB 200 DYEQIMEAWSDKGLYV 216

RESULT 15

rf2 nuclear restorer protein - maize
N/Alternate names: aldehyde dehydrogenase homolog
C/Species: Zea mays (maize)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: T03983
R/Cui, X.; Wise, R.P.; Schnable, P.S.
Science 272, 1334-1336, 1996
A/Title: The rf2 nuclear restorer gene of male-sterile T-cytoplasma maize.
A/Reference number: Z15177; MUID:96243131; PMID:8650543
A/Accession: T03983
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-549 <CUI>
A/Cross-references: EMBL:U43082; NID:g1421729; PIDN:AA049371.1; PID:g1421730
A/Experimental source: strain B73
C/Genetics:
A/Map position: 9
A/Note: rf2
C/Suprafamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
F:107-370/Domain: aldehyde dehydrogenase homology <ALD>

Query Match 41.6%; Score 49.5; DB 2; Length 549;
Best Local Similarity 58.8%; Pred. No. 1.1;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLLEAWSNDPTP 18
DB 147 NDHALALFTW-DNGKPY 162

Search completed: November 13, 2003, 09:50:21
Job time: 15.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHSLTLEAWSNDTPYLHD 21

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	108	90.8	1590	2	Q59983	O59983 streptococ
2	108	90.8	1590	2	O55263	O55263 streptococ
3	78	65.5	1338	2	Q9WJ34	Q9WJ34 streptococ
4	76	63.9	1016	2	Q9LJC7	Q9LJC7 leucococ
5	74	62.2	1477	2	Q9LJ66	Q9LJ66 leucococ
6	74	62.2	1508	2	Q9EZH5	Q9EZH5 leucococ
7	74	62.2	1508	2	O52224	O52224 leucococ
8	73	61.3	1575	2	Q9LJC3	Q9LJC3 streptococ
9	73	61.3	1577	2	O54178	O54178 streptococ
10	71	59.7	1527	2	Q9ZAR4	Q9ZAR4 leucococ
11	71	59.7	1527	2	Q8KRE1	Q8KRE1 leucococ
12	69	58.0	2835	2	Q8G9Q2	Q8G9Q2 leucococ
13	66	55.5	1449	2	O68542	O68542 streptococ
14	66	55.5	1449	2	O55264	O55264 streptococ
15	63	52.9	1554	2	Q8KZL5	Q8KZL5 streptococ
16	63	52.9	1577	2	Q55265	Q55265 streptococ

17	62	52.1	1512	2	Q9WJ35	Q9WJ35 streptococ
18	60	50.4	2057	2	Q9R805	Q9R805 leucococ
19	58	48.7	1518	2	Q00600	Q00600 streptococ
20	57	47.9	1599	2	Q00599	Q00599 streptococ
21	55.5	46.6	282	10	Q8RW33	Q8RW33 allium cepa
22	52	43.7	4848	2	O07944	O07944 streptococ
23	51	42.9	424	10	Q94J77	Q94J77 arabidopsis
24	51	42.9	435	10	Q9LJ68	Q9LJ68 arabidopsis
25	50	42.0	305	10	Q8LDL1	Q8LDL1 arabidopsis
26	50	42.0	339	10	Q94979	Q94979 arabidopsis
27	50	42.0	347	10	Q9LY92	Q9LY92 arabidopsis
28	49.5	41.6	55	10	Q94FX9	Q94FX9 zea mays
29	49.5	41.6	55	10	Q93VW4	Q93VW4 zea mays
30	49.5	41.6	55	10	Q93WH7	Q93WH7 zea mays
31	49.5	41.6	56	10	Q94FX8	Q94FX8 zea mays
32	49.5	41.6	129	10	Q94FX6	Q94FX6 zea mays
33	49.5	41.6	547	10	Q8LS74	Q8LS74 sorghum bic
34	49.5	41.6	549	10	Q94G64	Q94G64 zea mays
35	49.5	41.6	549	10	Q43274	Q43274 zea mays
36	49	41.2	601	16	O05566	O05566 mycobacteri
37	49	41.2	797	16	Q98B32	Q98B32 rhizobium
38	48.5	40.8	283	5	Q9V1P7	Q9V1P7 drosophila
39	48.5	40.8	283	5	O8WSF4	O8WSF4 mycobacteri
40	48	40.3	490	16	O06579	O06579 mycobacteri
41	47.5	39.9	214	11	Q8C1P7	Q8C1P7 mus musculu
42	47.5	39.9	519	10	Q9ZUB6	Q9ZUB6 arabidopsis
43	47.5	39.9	534	10	O8S528	O8S528 arabidopsis
44	47.5	39.9	534	10	Q94C67	Q94C67 arabidopsis
45	39.9	39.9	549	10	Q9LIR2	Q9LIR2 oryza sativ

ALIGNMENTS

RESULT 1

PRELIMINARY; FRT; 1590 AA.
ID O59983
AC O59983;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5).
GN GTF1
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QWZ176;
RX MEDLINE=94146405; PubMed=8312602;
RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d Streptococcus sobrinus.";
RL DNA Seq. 4:19-27(1993).
DR EMBL: D13858; BAA02976.1; -;
DR InterPro: IPR002478; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 16.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW SIGNAL
FT SIGNAL 1
FT CHAIN 39 1590
FT SEQUENCE 1590 AA; 175955 MW; C3683A57CF3C2B08 CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
Best Local Similarity 85.7%; Pred. No. 5.5e-08;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANDHSLTLEAWSNDTPYLHD 21
Db 477 ANNHVSIVEAWSNDTPYLHD 497


```

RESULT 6
Q9EZH5 PRELIMINARY; PRT; 1508 AA.
ID Q9EZH5;
AC Q9EZH5;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robye J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87AE4F3A CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAWSNDPTPYLHD 21
Db 563 ANQHSLIEDMSHNDPEYVD 583

RESULT 7
OS2224 PRELIMINARY; PRT; 1508 AA.
ID OS2224;
AC OS2224;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Remaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Glycosyltransferase; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAWSNDPTPYLHD 21
Db 563 ANQHSLIEDMSHNDPEYVD 583

RESULT 8
Q9LCH3

```

```

ID Q9LCH3 PRELIMINARY; PRT; 1575 AA.
AC Q9LCH3;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RA Fujitawa T., Hoshino T., Oshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BAA95201.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2E543 CRC64;

Query Match 61.3%; Score 73; DB 2; Length 1575;
Best Local Similarity 77.8%; Pred. No. 0.018;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 HSLIEAWSNDPTPYLHD 21
Db 549 HSLIEAWSNDPTPYNKD 566

RESULT 9
OS4178 PRELIMINARY; PRT; 1577 AA.
ID OS4178;
AC OS4178; Q54247;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RA Wickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgg, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC43483.1;
DR EMBL; M89776; AAA26969.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase.

```

SQ SEQUENCE 1577 AA; 177805 MM; 5AB0328DC5E08D18 CRC64;

Query Match 61.3%; Score 73; DB 2; Length 1577;

Best Local Similarity 77.8%; Pred. No. 0.018; Mismatches 4; Indels 0; Gaps 0;

QY 4 HHSILEAMSDNDPTLYLD 21
Db 551 HHSILEAMSDNDPTLYLD 568

RESULT 10

ID Q9ZAR4 PRELIMINARY; PRT; 1527 AA.

AC Q9ZAR4;

DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Dextranucrase.

GN DEX.

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL B-512-F;

RA Bhargava R., Singh D.K.S.;

RT "Cloning and Molecular Characterization of Dextranucrase Gene from

Leuconostoc mesenteroides NRRL B-512F."

RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.

DR InterPro: IPR002479; Glyco_hydro_70.

DR Pfam: PF01473; Glyco_hydro_70.

DR Pfam: PF03324; Glyco_hydro_70; 1.

SQ SEQUENCE 1527 AA; 169705 MM; 1DFAFA237C743398 CRC64;

Query Match 59.7%; Score 71; DB 2; Length 1527;

Best Local Similarity 66.7%; Pred. No. 0.037; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDPTLYLD 21
Db 581 ANOHLSTLEDMWSDNDPTLYLD 601

RESULT 11

ID Q8KRE1 PRELIMINARY; PRT; 1527 AA.

AC Q8KRE1;

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Dextranucrase DsrD (EC 2.4.1.5).

GN DSRD.

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN [1]

RP SEQUENCE FROM N.A.

RC "Isolation and characterization of the dextranucrase DsrD of

Leuconostoc mesenteroides DsrD."

RT Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF017384; AAG61158.1;

DR InterPro: IPR002479; Glyco_hydro_70.

DR Pfam: PF01473; Glyco_hydro_70; 1.

DR Pfam: PF03324; Glyco_hydro_70; 1.

KM Transferrase; Glycosyltransferase.

SQ SEQUENCE 1527 AA; 169835 MM; F9D0DE220BD89668 CRC64;

Query Match 59.7%; Score 71; DB 2; Length 1527;

Best Local Similarity 66.7%; Pred. No. 0.037; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDPTLYLD 21
Db 581 ANOHLSTLEDMWSDNDPTLYLD 601

RESULT 12

ID Q8G9Q2 PRELIMINARY; PRT; 2835 AA.

AC Q8G9Q2;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Dextranucrase (EC 2.4.1.5) (Fragment).

GN DSR.

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=2231661; PubMed=12270834;

RA Bozonnet S., Dolis-Lafargue M., Fabre E., Pizzut S., Renaud-Simeon M.,

Mondan P., Willemot R.M.;

RT "Molecular characterization of DSR-E, an alpha-1,2 linkage

synthesizing dextranucrase with two catalytic domains."

RL EMBL: AF430204; CND22883.1;

DR Transferrase; Glycosyltransferase.

FT NON TER 1

SQ SEQUENCE 2835 AA; 313264 MM; D03262CDD735399D CRC64;

Query Match 58.0%; Score 69; DB 2; Length 2835;

Best Local Similarity 57.9%; Pred. No. 0.16; Mismatches 3; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDPTLYLD 19
Db 557 ANHSLILEDMWSDNDPTLYLD 575

RESULT 13

ID Q68542 PRELIMINARY; PRT; 1449 AA.

AC Q68542;

DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Glucosyltransferase N (Fragment).

GN GTRN.

OS Streptococcus salivarius.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1304;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=V1477;

RA Jaffe R.I.;

RT "Streptococcus salivarius V1477 gtfN."

RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF049609; AAC05156.1;

DR InterPro: IPR002479; Glyco_hydro_70.

DR InterPro: IPR003318; Glyco_hydro_70.

DR Pfam: PF01473; Glyco_hydro_70; 1.

DR Pfam: PF03324; Glyco_hydro_70; 1.

KM Transferrase.

FT NON TER 1449

SQ SEQUENCE 1449 AA; 159895 MM; 0700F6D748471BFB CRC64;

Query Match 55.5%; Score 66; DB 2; Length 1449;

Best Local Similarity 72.2%; Pred. No. 0.21; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLILEAMSDNDTPYLHD 21
 DB 539 HSLILEAMSHNDAYNED 556

RESULT 14

Q55264 PRELIMINARY; PRT; 1449 AA.
 AC Q55264;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RL Infect. Immun. 63:609-621 (1995).
 DR EMBL; J35495; AAC4142.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 55.5%; Score 66; DB 2; Length 1449;
 Best Local Similarity 72.2%; Pred. No. 0.21;
 Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLILEAMSDNDTPYLHD 21
 DB 539 HSLILEAMSHNDAYNED 556

RESULT 15

Q8KZL5 PRELIMINARY; PRT; 1554 AA.
 ID Q8KZL5;
 AC Q8KZL5;
 DT 01-OCT-2002 (TREMblrel. 22, Created)
 DT 01-OCT-2002 (TREMblrel. 22, last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21956884; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan.";
 RT Biochim. Biophys. Acta 1570:75-79 (2002).
 RL EMBL; AB089438; BAC07265.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase.

SQ SEQUENCE 1554 AA; 171676 MW; 6981BC1DAE24A73 CRC64;

Query Match 52.9%; Score 63; DB 2; Length 1554;
 Best Local Similarity 66.7%; Pred. No. 0.69;
 Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLILEAMSDNDTPYLHD 21
 DB 488 HSLILEAMSLNDQYNED 505

Search completed: November 13, 2003, 09:43:58
 Job time : 34.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 / Search time 11.7962 Seconds
(without alignments)

154.898 Million cell updates/sec

Title: US-09-290-049a-2

Sequence: 1 VPSYSPFRAHSEVQDLIA 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	1475	2 B3135	gtfB protein precursor
2	91	95.8	1375	2 UT0345	dextranucrase (EC
3	84	88.4	1592	2 A38175	glucosyltransferase
4	80	84.2	1508	2 T31098	probable dextran
5	79	83.2	1290	2 JC5473	dextranucrase (EC
6	73	76.8	1365	2 A41483	glucosyltransferase
7	65	68.4	1431	2 A45866	dextranucrase (EC
8	65	68.4	1449	2 T30857	glucosyltransferase
9	65	68.4	1449	2 T30552	glucosyltransferase
10	65	68.4	1518	2 A44811	glucosyltransferase
11	64	67.4	1599	2 S22737	glucosyltransferase
12	61	64.2	1577	2 T30858	glucosyltransferase
13	46	48.4	575	2 S46329	intermediate filament
14	46	48.4	631	2 H90599	conserved hypob
15	44	46.3	327	2 D86474	probable RING zinc
16	44	46.3	51	2 G82455	hypothetical prote
17	43	45.3	214	2 R8BSCA	regulatory protein
18	43	45.3	1260	2 A86323	protein F14D16.3 l
19	42	44.2	431	2 S52583	ctx protein - Erw
20	42	44.2	563	2 P81888	probable D-lactate
21	42	44.2	563	2 D81134	D-lactate dehydrog
22	42	44.2	644	2 A64877	exoribonuclease II
23	42	44.2	644	2 C90861	RNase II, mRNA deg
24	42	44.2	644	2 B85758	ribonuclease T2 ho
25	42	44.2	659	2 A64139	probable histone t
26	42	44.2	932	2 T40216	lysine decarboxyla
27	41.5	43.7	713	2 B64743	lysine decarboxyla
28	41.5	43.7	713	2 D85503	lysine decarboxyla
29	41.5	43.7	713	2 D90652	lysine decarboxyla

30	41	43.2	145	2 G69098	hypothetical prote
31	41	43.2	344	2 JH0511	inositol 2-dehydro
32	41	43.2	597	2 F82935	DNA polymerase III
33	41	43.2	644	2 AB0656	exoribonuclease II
34	41	43.2	934	2 AG1889	WD-40 repeat prote
35	41	43.2	1146	2 E70204	hypothetical prote
36	40.5	42.6	414	2 T06120	hypothetical prote
37	40	42.1	359	2 G82197	RsfA1/RsfA2 prote
38	40	42.1	431	2 B37802	ctx protein - Erw
39	40	42.1	436	2 H97270	thiamine biosynthe
40	40	42.1	534	2 S41735	cholesterol estera
41	40	42.1	548	2 S32615	triacylglycerol 11
42	40	42.1	549	2 JN0552	triacylglycerol 11
43	40	42.1	549	2 JN0551	triacylglycerol 11
44	40	42.1	570	2 A90015	PTS system, lactos
45	40	42.1	572	2 B28474	phosphotransferase

ALIGNMENTS

RESULT 1
B3135
gtfB protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C/Accession: B3135; A33128
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Accession: B3135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1475 <SHT>
A/Cross-references: GB:M17361; NID:G153639; PIDN:AA88588.1; PID:G153640
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
Submitted to the Protein Sequence Database, September 1990
A/Reference number: A33128
A/Accession: A33128
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-171,173-641, 'N', 643-1475 <SH2>
A/Experimental source: strain GS-5
C/Superfamily: cpl repeat homology
F/1096-1115/Domain: cpl repeat homology <CP1>
F/1224-1243/Domain: cpl repeat homology <CP2>
F/1289-1308/Domain: cpl repeat homology <CP3>
F/1354-1373/Domain: cpl repeat homology <CP4>
F/1419-1438/Domain: cpl repeat homology <CP5>

Query Match 100.0%; Score 95; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSPFRAHSEVQDLIA 19
DB 552 VPSYSPFRAHSEVQDLIA 570

RESULT 2
UT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N/Alternate names: sucrose 6-glucosyltransferase
C/Species: Streptococcus mutans
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C/Accession: UT0345; C33135
R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
A/Reference number: UT0345; MUID:89137980; PMID:2976010
A/Accession: UT0345
A/Molecule type: DNA

A;Residues: 1-1375 <UED>
 A;Experimental source: GS-5
 R;Shirota, T.; Ueda, S.; Kuramitsu, H.K.
 J;Bacteriol. 169, 4263-4270, 1987
 A;Title: Sequence analysis of the *glfB* gene from *Streptococcus mutans*.
 A;Reference number: A33135; MUID:87308013; PMID:3040685
 A;Accession: C33135
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-349 <SHI>
 A;Cross-references: GB:MT7361
 C;Genetics:
 A;Gene: *glfC*
 C;Function: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 A;Description: cpl repeat homology
 C;Superfamily: cpl repeat homology
 C;Keywords: duplication; glycosyltransferase; hexosyltransferase
 F;1-34/Domain: signal sequence #status predicted <SIG>
 F;35-1375/Product: glycosyltransferase #status predicted <MAT>
 F;1126-1145/Domain: cpl repeat homology <CP1>
 F;1253-1272/Domain: cpl repeat homology <CP2>
 F;1318-1337/Domain: cpl repeat homology <CP3>

Query Match 95.8%; Score 91; DB 2; Length 1375;
 Best Local Similarity 100.0%; Pred. No. 6e-07; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0;

OY 1 VPYSFIRAHDSVQDLI 18
 DB 578 VPYSFIRAHDSVQDLI 595

RESULT 3
 A;8175
 glucosyltransferase precursor - *Streptococcus sobrinus*
 C;Species: *Streptococcus sobrinus*
 C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
 C;Accession: A38175
 R;Abu, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
 J;Bacteriol. 173, 989-996, 1991
 A;Title: Peptide sequences for sucrose splitting and glucan binding within *Streptococcus*
 A;Reference number: A38175; MUID:91123227; PMID:1704006
 A;Accession: A38175
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1592 <ABO>
 A;Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g217033
 C;Superfamily: cpl repeat homology
 F;1093-1112/Domain: cpl repeat homology <CP1>
 F;1223-1241/Domain: cpl repeat homology <CP2>
 F;1287-1306/Domain: cpl repeat homology <CP3>
 F;1330-1351/Domain: cpl repeat homology <CP4>
 F;1352-1371/Domain: cpl repeat homology <CP5>
 F;1402-1420/Domain: cpl repeat homology <CP6>
 F;1465-1484/Domain: cpl repeat homology <CP7>
 F;1513-1532/Domain: cpl repeat homology <CP8>

Query Match 88.4%; Score 84; DB 2; Length 1592;
 Best Local Similarity 88.9%; Pred. No. 1.1e-05; Indels 0; Gaps 0;
 Matches 16; Conservative 1; Mismatches 1;

OY 1 VPYSFIRAHDSVQDLI 18
 DB 548 VPYSFIRAHDSVQDLI 565

RESULT 4
 T31098
 Probable dextranucrase (EC 2.4.1.5), extracellular - *Leuconostoc mesenteroides*
 C;Species: *Leuconostoc mesenteroides*
 C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C;Accession: T31098
 R;Monchois, V.; Remaud-Simeon, M.; Monnan, P.; Willemot, R.M.

FEWS Microbiol. Lett. 159, 307-315, 1998
 A;Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A;Reference number: Z20981; MUID:98164374; PMID:9503626
 A;Accession: T31098
 A;Status: preliminary; translated from GB/EMBL/DDBT
 A;Molecule type: DNA
 A;Residues: 1-1508 <MON>
 A;Cross-references: EMBL:AR030129; NID:g2766611; PID:g2766612; PIDN:AB95453.1
 A;Experimental source: strain NRRL B-1299
 C;Genetics:
 A;Gene: *dsrB*
 C;Function: produces dextran composed only of alpha(1-6) glucosidic bonds
 A;Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 84.2%; Score 80; DB 2; Length 1508;
 Best Local Similarity 73.7%; Pred. No. 5e-05; Indels 0; Gaps 0;
 Matches 14; Conservative 4; Mismatches 1;

OY 1 VPYSFIRAHDSVQDLI 19
 DB 634 IPNYSFVRAHDSVQTLA 652

RESULT 5
 JC5473
 dextranucrase (EC 2.4.1.5) - *Leuconostoc mesenteroides*
 C;Species: *Leuconostoc mesenteroides*
 C;Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C;Accession: JC5473
 R;Monchois, V.; Willemot, R.M.; Remaud-Simeon, M.; Croux, C.; Monnan, P.
 Gene 182, 23-32, 1996
 A;Title: Cloning and sequencing of a gene coding for a novel dextranucrase from *Leucon*
 A;Reference number: JC5473; MUID:97136686; PMID:8982063
 A;Accession: JC5473
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1290 <MON>
 A;Cross-references: GB:U38181
 C;Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose on
 C;Genetics:
 A;Gene: *dsrA*
 C;Keywords: glycosyltransferase; hexosyltransferase
 F;78-870/Domain: catalytic #status predicted <CAT>
 F;922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 83.2%; Score 79; DB 2; Length 1290;
 Best Local Similarity 83.3%; Pred. No. 6.2e-05; Indels 0; Gaps 0;
 Matches 15; Conservative 2; Mismatches 1;

OY 2 PYSFIRAHDSVQDLI 19
 DB 388 PYSFIRAHDSVQTLA 405

RESULT 6
 A41483
 glucosyltransferase (EC 2.4.1.-) *glfB* precursor - *Streptococcus sobrinus*
 C;Species: *Streptococcus sobrinus*
 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C;Accession: A41483
 R;Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
 Infect. Immun. 58, 2452-2458, 1990
 A;Title: Analysis of the *Streptococcus downii glfB* gene, which specifies a glucosyltra
 A;Reference number: A41483; MUID:90316665; PMID:2142479
 A;Accession: A41483
 A;Molecule type: DNA
 A;Residues: 1-1365 <GIL>
 A;Cross-references: GB:M30943; NID:g153652; PIDN:AAA6898.1; PID:g153653
 C;Genetics:
 A;Gene: *glfB*
 C;Superfamily: cpl repeat homology
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 76.8%; Score 73; DB 2; Length 1365;
 Best Local Similarity 78.9%; Pred. No. 0.00071;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSVQDLIA 19
 ||:|||||||
 DB 537 VPMYFIRAHDSVQTRIA 555

RESULT 7

dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
 C/Species: Streptococcus mutans
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C/Accession: A45866
 R/Honda, O.; Kato, C.; Kuramitsu, H.K.
 U. Gen. Microbiol. 136, 2089-2105, 1990
 A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
 A/Reference number: A45866; MVID:91100958; PMID:2148600
 A/Accession: A45866
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1431 <HON>
 A/Cross-references: GB:M29296
 C/Superfamily: cpl repeat homology
 C/Keywords: glycosyltransferase; hexosyltransferase
 F:181-201/Domain: cpl repeat homology <CP1>
 F:1127-1146/Domain: cpl repeat homology <CP2>
 F:1192-1211/Domain: cpl repeat homology <CP3>
 F:1257-1276/Domain: cpl repeat homology <CP4>
 F:1277-1297/Domain: cpl repeat homology <CP5>
 F:1321-1340/Domain: cpl repeat homology <CP6>
 F:1341-1361/Domain: cpl repeat homology <CP7>
 F:1385-1406/Domain: cpl repeat homology <CP7>

Query Match 68.4%; Score 65; DB 2; Length 1431;
 Best Local Similarity 76.5%; Pred. No. 0.017;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 19
 :|||
 DB 576 NYFIRAHDSVQTVIA 592

RESULT 8

glucosyltransferase - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C/Accession: T30857
 R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
 Infect. Immun. 63, 609-621, 1995
 A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pro
 A/Reference number: Z20909; MVID:95122197; PMID:7822030
 A/Accession: T30857
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1449 <SIM>
 A/Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAQ41412.1
 C/Genetics:
 A/Gene: gtfL

Query Match 68.4%; Score 65; DB 2; Length 1449;
 Best Local Similarity 68.8%; Pred. No. 0.018;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 18
 :|||
 DB 609 NYAFVRAHDSVQSII 624

RESULT 9

T30552
 glucosyltransferase N - Streptococcus salivarius (fragment)

C/Species: Streptococcus salivarius
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C/Accession: T30552
 R/Jaffe, R.I.
 submitted to the EMBL Data Library, February 1998
 A/Description: Streptococcus salivarius V1477 gtfN.
 A/Reference number: Z20854
 A/Accession: T30552
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1449 <JAF>
 A/Cross-references: EMBL:RF049609; NID:g2935545; PID:g2935546; PIDN:AAQ05156.1
 C/Genetics:
 A/Gene: gtfN

Query Match 68.4%; Score 65; DB 2; Length 1449;
 Best Local Similarity 68.8%; Pred. No. 0.018;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 18
 :|||
 DB 609 NYAFVRAHDSVQSII 624

RESULT 10

glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
 C/Accession: A44811; S22726; S28809
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 U. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
 A/Reference number: A44811; MVID:92148377; PMID:1858391
 A/Accession: A44811
 A/Molecule type: DNA
 A/Residues: 1-1518 <GIF>
 A/Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAAT7900.1; PID:g47527
 A/Note: sequence extracted from NCBI backbone (NCBI:81050, NCBI:81052)
 C/Genetics:
 A/Gene: gtfY
 C/Superfamily: cpl repeat homology
 C/Keywords: glycosyltransferase; hexosyltransferase
 F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 68.4%; Score 65; DB 2; Length 1518;
 Best Local Similarity 70.6%; Pred. No. 0.019;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 19
 :|||
 DB 604 NYFIRAHDSVQDLIA 620

RESULT 11

glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
 C/Accession: S22737; S28810; B44811; S22727
 R/Jacques, N. the EMBL Data Library, March 1992
 submitted to the EMBL Data Library, March 1992
 A/Reference number: S22737
 A/Accession: S22737
 A/Molecule type: DNA
 A/Residues: 1-1599 <JAC>
 A/Cross-references: EMBL:Z11872; NID:g47530; PIDN:CAAT7898.1; PID:g47531
 A/Experimental source: ATCC 25975
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 U. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

PI Nichols SE;
 XX WPI: 2002-414332/44.
 DR N-PSDB; ABK52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 21-25; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.

XX Sequence 1475 AA;
 SQ
 Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPSYSTRAHDSVQDLIA 19
 DB 552 VPSYSTRAHDSVQDLIA 570

RESULT 2
 AAU98030 ID AAU98030 standard; Protein; 1475 AA.
 XX AAU98030;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB mutant I448V.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX

OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT
 FT US2002031826-A1.
 XX
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0462711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 PI Nichols SE;
 XX WPI: 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX

SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDLIA 19
 |||||
 DB 552 VPSYSFIRAHSEVQDLIA 570

RESULT 3
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX AAU98031;
 AC
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutein.
 KM
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 457 /note="Wild-type Asp substituted by Asn"
 FT
 XX
 XX US2002031826-A1.
 PN
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 XX
 XX Nichols SE;
 PI
 DR WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Claim 36; Page -; 44p; English.
 PS
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes PI or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll a/b binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDLIA 19
 |||||
 DB 552 VPSYSFIRAHSEVQDLIA 570

RESULT 4
 AAU98033
 ID AAU98033 standard; Protein; 1475 AA.
 XX AAU98033;
 AC
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant K1014T.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutein.
 KM
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1014 /note="Wild-type Lys substituted by Thr"
 FT
 XX
 XX US2002031826-A1.
 PN
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 XX

CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVODLIA 19
Db 552 VPSYSFIRAHSEVODLIA 570

RESULT 6
AAU98039 standard; Protein; 1475 AA.

AC AAU98039;

DT 27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan, starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutein.

Streptococcus mutans.
Synthetic.

Key Location/Qualifiers

FT Misc-difference 169..171

FT /note= "Wild-type Tyr-Tyr-Tyr substituted by
Ala-Ala-Ala"

PN US2002031826-A1.

PD 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substitutes for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions -

PS Claim 36, Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilising the glucan produced by GTF, which utilises
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention but
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVODLIA 19
Db 552 VPSYSFIRAHSEVODLIA 570

RESULT 7
AAU98040 standard; Protein; 1475 AA.

AC AAU98040;

DT 27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant K779Q.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan, starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutein.

Streptococcus mutans.
Synthetic.

Key Location/Qualifiers

FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"

PN US2002031826-A1.

PD 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR MPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, 1589E, N471D, N471D/1589D, and
 CC N471D/1589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 XX
 QY Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 VPSYSFIRAHDSVQDLIA 19
 552 VPSYSFIRAHDSVQDLIA 570
 XX
 RESULT 8
 AAU79285
 ID AAU79285 standard; Protein; 1017 AA.
 XX
 AC AAU79285;
 XX
 DT 13-AUG-2002 (first entry)

XX
 DE Streptococcus mutans monoclonal antibody-related protein #2.
 XX
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-carries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 EN JP2002114709-A.
 XX
 PD 16-APR-2002.
 XX
 PF 04-OCT-2000; 2000JP-0304889.
 XX
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PA (UYNI-) UNIV NIPPON.
 XX
 DR MPI; 2002-448101/48.
 XX
 PT Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 PS Claim 4; Page 17-19; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP16 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 XX
 SQ Sequence 1017 AA;
 XX
 QY Query Match 95.8%; Score 91; DB 23; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 VPSYSFIRAHDSVQDLI 18
 518 VPSYSFIRAHDSVQDLI 535
 XX
 RESULT 9
 AAU98028
 ID AAU98028 standard; Protein; 1375 AA.
 XX
 AC AAU98028;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 XX
 EN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0428704.
 PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI: 2002-414332/44.
 XX N-PSDB; ABK52939.
 DR
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 30-33; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779G/K1014T,
 CC Y169A/Y170A/Y171A, and K779G or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFc.
 CC
 XX
 SQ Sequence 1375 AA;
 QY
 Db Query Match 95.8%; Score 91; DB 23; Length 1375;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLI 18
 |||||
 Db 578 VPSYSFIRAHDSVQDLI 595

RESULT 10
 AAU79288
 ID AAU79288 standard; Protein; 1375 AA.
 AC AAU79288;
 XX
 DT 13-AUG-2002 (first entry)

XX
 DE Streptococcus mutans monoclonal antibody-related protein #5.
 XX
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-caries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 PN JP2002114709-A.
 XX
 PD 16-APR-2002.
 XX
 PF 04-OCT-2000; 2000JP-0304889.
 XX
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PA (UYN1-) UNIV NIPPON.
 XX
 DR WPI: 2002-448101/48.
 XX
 PT Anti-caries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 PS Disclosure; Page 22-25; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-caries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitory activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 XX
 SQ Sequence 1375 AA;
 QY
 Db Query Match 95.8%; Score 91; DB 23; Length 1375;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLI 18
 |||||
 Db 578 VPSYSFIRAHDSVQDLI 595

RESULT 11
 AAU79284
 ID AAU79284 standard; Protein; 1476 AA.
 AC AAU79284;
 XX
 DT 13-AUG-2002 (first entry)

DE Streptococcus mutans monoclonal antibody-related protein #1.
 XX
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-caries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 PN JP2002114709-A.
 XX
 PD 16-APR-2002.
 XX
 PF 04-OCT-2000; 2000JP-0304889.
 XX
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PA (UYN1-) UNIV NIPPON.

DR WPI; 2002-448101/48.

XX Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl

PT transferase-B (GTF-B) of Streptococcus mutans -

PS Claim 3; Page 13-16; 28pp; Japanese.

XX The invention relates to a monoclonal antibody against dental caries and

CC an anti-carries agent composed of a monoclonal antibody produced by

CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)

CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having

CC inhibitive activity against water insoluble glucan synthetase of glucosyl

CC transferase-B. The monoclonal antibody specifically inhibits water

CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl

CC transferase-B and is used in the immunotherapy of dental caries. This

CC sequence represents a Streptococcus mutans monoclonal antibody-related

CC protein.

XX Sequence 1476 AA;

SQ

Query Match 95.8%; Score 91; DB 23; Length 1476;

Best Local Similarity 100.0%; Pred. No. 1.6e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVODLI 18

DB 552 VPSYSFIRAHDSVODLI 569

RESULT 12

AAU98032 standard; Protein; 1475 AA.

ID AAU98032;

XX AAU98032;

AC

XX 27-AUG-2002 (first entry)

DT

XX

DB S. mutans glucosyltransferase GTFB mutant D567T.

XX

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

XX

XX Streptococcus mutans.

OS Synthetic.

XX

XX Key Location/Qualifiers

FH Misc-difference 567 /note="Wild-type Asp substituted by Thr"

FT

XX US2002031826-A1.

PN

XX 14-MAR-2002.

PD

XX 19-DEC-2000; 2000US-0740274.

PF

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

PR

XX (NICH/) NICHOLS S E.

PA

XX Nichols SE;

PI

XX WPI; 2002-414332/44.

DR

XX Glucosyltransferase B or D protein useful for producing a glucan useful

PT as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -; 44pp; English.

PS

XX The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from 1448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, D448V/D457N/D567T/K779C/K1014T,

CC Y159A/Y170A/Y171A, and K779C or a GTF D polypeptide having

CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the

CC vector, a seed or tuber from the transgenic plant, a paper transformed

CC coating composition comprising a glucan produced in a plant where

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or or glucan and starch where

CC the glucan is produced in the amyloplast and/or vacuole or a maize line

CC deficient in starch biosynthesis, transformed with a gene encoding a

CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful

CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing

CC conditions to produce a regenerated plant and inducing expression of the

CC polynucleotide for a time sufficient to produce the glucan in the

CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC is produced in the amyloplast of potato or the vacuole of sugar beet.

CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture, unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper

CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention but

CC Note: The present sequence is not shown in the specification but

CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.

XX

SQ Sequence 1475 AA;

Query Match 92.6%; Score 88; DB 23; Length 1475;

Best Local Similarity 94.7%; Pred. No. 5.8e-07;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVODLI 19

DB 552 VPSYSFIRAHDSVODLI 570

RESULT 13

AAU98034 standard; Protein; 1475 AA.

ID AAU98034;

XX AAU98034;

AC

XX 27-AUG-2002 (first entry)

DT

XX

DB S. mutans glucosyltransferase GTFB mutant D457N/D567T.

XX

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

XX

XX Streptococcus mutans.

OS Synthetic.

XX

XX Key Location/Qualifiers

FH

FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -, 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D571K, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D571K/K790Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AA098C27
 CC and the information in claim 36.
 CC Sequence 1475 AA;
 SQ

Query Match 92.6%; Score 88; DB 23; Length 1475;
 Best Local Similarity 94.7%; Pred. No. 5.8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPSSYFIRAHDSFVODLII 19
 DB 552 VPSSYFIRAHDSFVODLII 570
 RESULT 14
 AAU98036
 ID AAU98036 standard; Protein; 1475 AA.
 XX
 AC AAU98036;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT FT
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -, 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D571K, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D571K/K790Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC

thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing the glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch. CC and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. CC The present sequence represents a GTF mutant of the invention. CC Note: The present sequence is not shown in the specification but CC was created by the indexer using the GTFB sequence appearing as AAU98027 CC and the information in claim 36.

Sequence 1475 AA:

Query Match 92.6%; Score 88; DB 23; Length 1475;
Best Local Similarity 94.7%; Pred. No. 5,8e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VPSYSTRAHDSVQDLIA 19
|||||
552 VPSYSTRAHDSVQDLIA 570

RESULT 15

AAU98037 standard; Protein; 1475 AA.

AAU98037;

27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;

coating composition; glucan; starch; latex; thermoplastic molecule;

amyloplast; vacuole; paper manufacture; mutant; mutain.

Streptococcus mutans.

Synthetic.

Key Location/Qualifiers

Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0704274.

11-DEC-1998; 98US-0210361.

07-JUN-1995; 95US-0478704.

07-JUN-1995; 95US-0482711.

07-JUN-1995; 95US-0485243.

16-JAN-1998; 98US-0007999.

16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

DR WPI; 2002-414332/44.

Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions

Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T/D571K, D457N/D571K, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. CC The present sequence represents a GTFB mutant of the invention. CC Note: The present sequence is not shown in the specification but CC was created by the indexer using the GTFB sequence appearing as AAU98027 CC and the information in claim 36.

Sequence 1475 AA:

Query Match 92.6%; Score 88; DB 23; Length 1475;
Best Local Similarity 94.7%; Pred. No. 5,8e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VPSYSTRAHDSVQDLIA 19
|||||
552 VPSYSTRAHDSVQDLIA 570

Search completed: November 13, 2003, 09:38:25
Job time : 34.8463 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 12.365 Seconds
(Without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-2

Sequence: 1 VPSYSFIRAHDSVQDLIA 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/6C_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	1475	US-09-007-999-2	Sequence 2, Appli
2	95	100.0	1475	US-09-210-361-2	Sequence 2, Appli
3	95	100.0	1475	US-09-740-274-2	Sequence 2, Appli
4	91	95.8	1375	US-09-210-361-4	Sequence 4, Appli
5	91	95.8	1375	US-09-740-274-4	Sequence 4, Appli
6	80	84.2	523	US-09-604-957-5	Sequence 5, Appli
7	67	70.5	584	US-09-604-957-6	Sequence 6, Appli
8	66	69.5	535	US-09-604-957-7	Sequence 7, Appli
9	66	69.5	1278	US-09-604-957-3	Sequence 3, Appli
10	66	69.5	2057	US-09-499-203-2	Sequence 2, Appli
11	65	68.4	545	US-09-604-957-4	Sequence 4, Appli
12	65	68.4	1430	US-09-008-117-2	Sequence 2, Appli
13	65	68.4	1430	US-09-210-361-6	Sequence 6, Appli
14	65	68.4	1430	US-09-740-274-6	Sequence 6, Appli
15	61	64.2	1577	US-08-793-824-2	Sequence 2, Appli
16	43	45.3	75	US-09-107-532A-7020	Sequence 7020, Ap
17	42	44.2	565	US-09-107-532A-4217	Sequence 4217, Ap
18	42	44.2	616	US-09-268-347-38	Sequence 38, Appli
19	42	44.2	659	US-09-268-347-46	Sequence 46, Appli
20	42	44.2	660	US-09-268-347-45	Sequence 45, Appli
21	42	44.2	2627	US-08-751-189-3	Sequence 3, Appli
22	42	44.2	2627	US-09-060-836-3	Sequence 3, Appli
23	42	44.2	2627	US-09-184-445-3	Sequence 3, Appli
24	41.5	43.7	713	US-08-849-212-4	Sequence 4, Appli
25	40	42.1	431	US-07-783-705A-2	Sequence 2, Appli
26	39	41.1	309	US-09-345-473E-37	Sequence 37, Appli
27	38.5	40.5	237	US-08-578-709-11	Sequence 11, Appli

28	38.5	40.5	345	3	US-09-231-529-1	Sequence 1, Appli
29	38.5	40.5	345	3	US-08-977-816-1	Sequence 1, Appli
30	38.5	40.5	350	4	US-09-533-494A-17	Sequence 17, Appli
31	38.5	40.5	417	4	US-09-533-494A-13	Sequence 13, Appli
32	38.5	40.5	500	2	US-08-578-709-15	Sequence 15, Appli
33	38	40.0	184	1	US-08-353-550-2	Sequence 2, Appli
34	38	40.0	184	1	US-08-551-687-2	Sequence 2, Appli
35	38	40.0	220	4	US-09-252-991A-18490	Sequence 18490, A
36	38	40.0	331	4	US-09-252-991A-18759	Sequence 18759, A
37	38	40.0	333	4	US-09-252-991A-33113	Sequence 33113, A
38	38	40.0	325	4	US-09-107-532A-6786	Sequence 6786, Ap
39	38	40.0	387	4	US-09-134-001C-3206	Sequence 3206, Ap
40	38	40.0	389	4	US-09-107-532A-6456	Sequence 6456, Ap
41	38	40.0	428	1	US-08-353-550-1	Sequence 1, Appli
42	38	40.0	428	1	US-08-551-687-1	Sequence 1, Appli
43	38	40.0	431	3	US-08-807-342B-5	Sequence 5, Appli
44	38	40.0	469	1	US-08-353-550-6	Sequence 6, Appli
45	38	40.0	469	2	US-08-551-687-6	Sequence 6, Appli

ALIGNMENTS

```

RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: Latexes in Paper Manufacture
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match      100.0%; Score 95; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 VPSYSFIRAHDSVQDLIA 19
Db      552 VPSYSFIRAHDSVQDLIA 570

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: Latexes in Paper Manufacture
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match          100.0%; Score 95; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLIA 19
Db      552 VPSYSFIRAHDSVQDLIA 570

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match          100.0%; Score 95; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLIA 19
Db      552 VPSYSFIRAHDSVQDLIA 570

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          95.8%; Score 91; DB 3; Length 1375;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLI 18
Db      578 VPSYSFIRAHDSVQDLI 595

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          95.8%; Score 91; DB 4; Length 1375;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLI 18
Db      578 VPSYSFIRAHDSVQDLI 595

RESULT 6
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
```

```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-604-957-5

Query Match          84.2%; Score 80; DB 4; Length 523;
Best Local Similarity 73.7%; Pred. No. 2.3e-06;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLI 19
Db      146 IPNYSFVRADHSDVOTVIA 164

RESULT 7
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-604-957-6

Query Match          70.5%; Score 67; DB 4; Length 584;
Best Local Similarity 57.9%; Pred. No. 0.00062;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLI 19
Db      167 IPNYSFVRADHDADDPIS 185

RESULT 8
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
```

```

; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
; US-09-604-957-7

Query Match          69.5%; Score 66; DB 4; Length 535;
Best Local Similarity 61.1%; Pred. No. 0.00084;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLI 18
Db      144 IPNYSFVRADHSDSDQI 161

RESULT 9
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
; US-09-604-957-3

Query Match          69.5%; Score 66; DB 4; Length 1278;
Best Local Similarity 61.1%; Pred. No. 0.0024;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLI 18
Db      620 IPNYSFVRADHSDSDQI 637

RESULT 10
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: CUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-499-203-2

Query Match          69.5%; Score 66; DB 4; Length 2057;
```

Best Local Similarity 61.1%; Pred. No. 0.0043;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSFIRAHSEVODLIA 18
:|||||:|
DB 757 INYSFVRAHDVADADPEI 774

RESULT 11
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHMOU, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match 68.4%; Score 65; DB 4; Length 545;
Best Local Similarity 76.5%; Pred. No. 0.0013; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHSEVODLIA 19
:|||||:|
DB 156 NYIFIRAHSEVQTVIA 172

RESULT 12
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 68.4%; Score 65; DB 3; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.0042; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHSEVODLIA 19
:|||||:|
DB 576 NYIFIRAHSEVQTVIA 592

RESULT 13
US-09-210-361-6
; Sequence 6, Application US/09210361

; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 68.4%; Score 65; DB 3; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.0042; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHSEVODLIA 19
:|||||:|
DB 576 NYIFIRAHSEVQTVIA 592

RESULT 14
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 68.4%; Score 65; DB 4; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.0042; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHDESEVODLIA 19
 Db 576 NYIFIRAHDESEVQAVIA 592

RESULT 15

US-08-793-824-2
 ; Sequence 2, Application US/08793824
 ; Patent No. 5981838
 ; GENERAL INFORMATION:
 ; APPLICANT: Simpson, Christine Lynn
 ; APPLICANT: Giffard, Philip Morrison
 ; APPLICANT: Jacques, Nicholas Anthony
 ; TITLE OF INVENTION: Genetic Manipulation of Plants to
 ; TITLE OF INVENTION: Increase Stored Carbohydrates
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Griffith Hack & Co
 ; STREET: Level 8, 168 Walker Street
 ; CITY: No. 5981838th Sydney
 ; STATE: New South Wales
 ; COUNTRY: Australia
 ; ZIP: 2060
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/793,824
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: AU PM7643
 ; FILING DATE: 24-AUG-1994
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 61 2 9957 5944
 ; TELEFAX: 61 2 957 6288
 ; TELEX: 26547
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1577 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus salivarius
 ; US-08-793-824-2

Query Match 64.2%; Score 61; DB 2; Length 1577;
 Best Local Similarity 64.7%; Pred. No. 0.025;
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 3 SYSFIRAHDESEVODLIA 19
 Db 661 NYIFIRAHDESEVQAVIA 677

Search completed: November 13, 2003, 09:47:57
 Job time : 12.3365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 : Search time 24.8531 Seconds

(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-2

Perfect score: 95

Sequence: 1 VPSSYFIRAHDSFVQDLIA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	1475	9	US-09-740-274-2
2	91	95.8	1375	9	US-09-740-274-4
3	80	84.2	522	10	US-09-995-749A-11
4	66	69.5	535	10	US-09-995-749A-12
5	66	69.5	584	10	US-09-995-749A-12
6	66	69.5	1781	10	US-09-995-749A-2
7	65	68.4	545	10	US-09-995-749A-10
8	65	68.4	1430	9	US-09-740-274-6
9	42	44.2	115	11	US-09-764-891-5098
10	42	44.2	364	15	US-10-156-761-7564
11	41	43.2	170	9	US-09-775-925-31
12	41	43.2	458	15	US-10-103-813-362
13	41	43.2	552	10	US-09-955-732-15
14	41	43.2	552	10	US-09-955-732-15
15	41	43.2	662	10	US-09-738-626-6490

16	40.5	42.6	296	14	US-10-080-960-19	Sequence 19, Appl
17	40	42.1	170	9	US-08-775-925-82	Sequence 32, Appl
18	40	42.1	241	9	US-09-775-925-8	Sequence 8, Appl
19	40	42.1	484	12	US-10-181-590-9	Sequence 9, Appl
20	40	42.1	509	9	US-09-775-925-6	Sequence 6, Appl
21	40	42.1	509	10	US-09-955-732-14	Sequence 14, Appl
22	40	42.1	547	11	US-09-943-857-2	Sequence 2, Appl
23	40	42.1	547	11	US-09-943-857-4	Sequence 4, Appl
24	40	42.1	547	11	US-09-943-857-6	Sequence 6, Appl
25	40	42.1	572	9	US-09-815-242-5626	Sequence 5626, Ap
26	40	42.1	572	9	US-09-815-242-12247	Sequence 12247, A
27	40	42.1	573	9	US-09-815-242-12921	Sequence 12921, A
28	39	41.1	67	15	US-10-007-280A-214	Sequence 214, Ap
29	39	41.1	309	10	US-09-862-027-37	Sequence 37, Appl
30	39	41.1	405	15	US-10-156-761-9148	Sequence 9148, Ap
31	39	41.1	489	10	US-09-738-626-3805	Sequence 3805, Ap
32	39	41.1	557	11	US-09-934-455-326	Sequence 326, Appl
33	39	41.1	779	12	US-10-353-929-49	Sequence 82, Appl
34	39	41.1	1383	15	US-10-021-955-82	Sequence 86, Appl
35	39	41.1	1383	15	US-10-021-955-86	Sequence 87, Appl
36	39	41.1	1383	15	US-10-021-955-87	Sequence 79, Appl
37	39	41.1	1389	15	US-10-021-955-79	Sequence 45372, A
38	38	40.0	166	9	US-09-864-761-45372	Sequence 5731, Ap
39	38	40.0	312	10	US-09-738-626-5731	Sequence 12045, A
40	38	40.0	421	15	US-10-156-761-12045	Sequence 3363, Ap
41	38	40.0	574	9	US-09-815-242-12045	Sequence 2, Appl
42	38	40.0	831	15	US-10-128-714-3363	Sequence 7, Appl
43	38	40.0	1025	15	US-10-055-475-2	Sequence 9, Appl
44	38	40.0	1025	15	US-10-055-475-7	
45	38	40.0	1025	15	US-10-055-475-9	

ALIGNMENTS

RESULT 1

US-09-740-274-2

Sequence 2, Application US/09740274

Patent No. US20020031826A1

GENERAL INFORMATION:

APPLICANT: Nichols, Scott E.

TITLE OF INVENTION: Glucan-containing Compositions and Paper

FILE REFERENCE: 0357CRD

CURRENT APPLICATION NUMBER: US/09/740,274

CURRENT FILING DATE: 2000-12-19

PRIOR APPLICATION NUMBER: 09/210,361

PRIOR FILING DATE: 1998-12-11

PRIOR APPLICATION NUMBER: 09/007,999

PRIOR FILING DATE: 1998-01-16

PRIOR APPLICATION NUMBER: 08/478,704

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: 09/009,620

PRIOR FILING DATE: 1998-01-20

PRIOR APPLICATION NUMBER: 08/485,243

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: 09/008,172

PRIOR FILING DATE: 1998-01-16

PRIOR APPLICATION NUMBER: 08/482,711

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 1475

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-740-274-2

Query Match

Best Local Similarity 100.0%; Score 95; DB 9; Length 1475;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VPSSYFIRAHDSFVQDLIA 19

|||||

|||||

Db 552 VPSYSTRADSEVODLI 570

RESULT 2
US-09-740-274-4
Sequence 4, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: NICHOLS, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 95.8%; Score 91; DB 9; Length 1375;
Best Local Similarity 100.0%; Pred. No. 7.9e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSTRADSEVODLI 18
Db 578 VPSYSTRADSEVODLI 595

RESULT 3
US-09-995-749A-11
Sequence 11, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: BO43388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 522
TYPE: PRT
ORGANISM: leuconostoc mesenteroides
US-09-995-749A-11

Query Match 84.2%; Score 80; DB 10; Length 522;
Best Local Similarity 73.7%; Pred. No. 2.1e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSTRADSEVODLI 19
Db 146 IPNYSFVRADSEVODTIA 164

RESULT 4
US-09-995-749A-13
Sequence 13, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: BO43388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 535
TYPE: PRT
ORGANISM: lactobacillus reuteri
US-09-995-749A-13

Query Match 69.5%; Score 66; DB 10; Length 535;
Best Local Similarity 61.1%; Pred. No. 0.0057;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSTRADSEVODLI 18
Db 144 IPNYSFVRADHNNSDQI 161

RESULT 5
US-09-995-749A-12
Sequence 12, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: BO43388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 584
TYPE: PRT
ORGANISM: leuconostoc mesenteroides
US-09-995-749A-12

Query Match 69.5%; Score 66; DB 10; Length 584;
Best Local Similarity 61.1%; Pred. No. 0.0062;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSTRADSEVODLI 18
Db 167 IPNYSFVRADHNDQDPI 184

```

APPLICANT: Nicholas, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US/09/740,274
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match          68.4%; Score 65; DB 9; Length 1430;
Best Local Similarity 76.5%; Pred.No. 0.026;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

CY      3 SYSFIKRAHSEVQDLIA 19
       :|:|||||::|||
Db      576 NYIFIRAHSEVQTIVA 592

RESULT 9
US-09-764-891-5098
Sequence 5098, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5098
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (3)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (9)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (109)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (112)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-5098

Query Match          44.2%; Score 42; DB 11; Length 115;
Best Local Similarity 38.9%; Pred.No. 14;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

CY      2 PSYSFIKRAHSEVQDLIA 19

```

Db 63 PAAVFRHRAQRCQVVA 80

RESULT 10

US-10-156-761-7564
; Sequence 7564, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIRAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHITRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 7564
LENGTH: 364
TYPE: PRT
ORGANISM: Streptomyces avermectilis
US-10-156-761-7564

Query Match 44.2%; Score 42; DB 15; Length 364;
Best Local Similarity 44.4%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PSYSFIRAHDSFVODLIA 19
Db 15 PEIKGRHDEDEALIELLA 32

RESULT 11

US-09-775-925-31
; Sequence 31, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:

APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
FILE OF INVENTION: PHOSPHATASES
FILE REFERENCE: 200125.420
CURRENT APPLICATION NUMBER: US/09/775,925
CURRENT FILING DATE: 2001-02-01
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-775-925-31

Query Match 43.2%; Score 41; DB 9; Length 170;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSFVODLIA 19
Db 72 AYHNIRVYDEETDILIA 88

RESULT 12
US-10-103-313-362
; Sequence 362, Application US/10103313
; Publication No. US20030082758A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 362
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (304)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc.feature
; LOCATION: (308)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc.feature
; LOCATION: (377)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-362

Query Match 43.2%; Score 41; DB 15; Length 458;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSFVODLIA 19
Db 41 AYHNIRVYDEETDILIA 57

RESULT 13

US-09-775-925-2
; Sequence 2, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:

APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
FILE OF INVENTION: PHOSPHATASES
FILE REFERENCE: 200125.420
CURRENT APPLICATION NUMBER: US/09/775,925
CURRENT FILING DATE: 2001-02-01
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 552
TYPE: PRT
ORGANISM: Homo sapiens
US-09-775-925-2

Query Match 43.2%; Score 41; DB 9; Length 552;
Best Local Similarity 47.1%; Pred. No. 12e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSFVODLIA 19
Db 214 AYHNIRVYDEETDILIA 230

RESULT 14
US-09-955-732-15
; Sequence 15, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:

APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.433
CURRENT APPLICATION NUMBER: US/09/955,732

```

; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 552
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-955-732-15

```

```

Query Match      43.2% Score 41; DB 10; Length 552;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches      8; Conservative      3; Mismatches      6; Indels      0; Gaps      0;

```

```

QY      3 SYSFIRAHSEVODLIA 19
      :|||:|||||:
Db      214 AYHNRVYDEETDILA 230

```

```

RESULT 15
US-09-738-626-6490
; Sequence 6490, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6490
; LENGTH: 662
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6490

```

```

Query Match      43.2% Score 41; DB 10; Length 662;
Best Local Similarity 53.3%; Pred. No. 1.5e+02;
Matches      8; Conservative      2; Mismatches      5; Indels      0; Gaps      0;

```

```

QY      5 SFIRAHSEVODLIA 19
      :|||:|||||:
Db      415 SIEKLHDSPAQDLVA 429

```

Search completed: November 13, 2003, 10:23:00
Job time : 25.8531 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 6.8436 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-2

Perfect score: 95
Sequence: 1 VPSYSFIRAHSEVQDLIA 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	95.8	1455	1	GTFC_STRMU
2	91	95.8	1476	1	GTFC_STRMU
3	86	90.5	1597	1	GTFC_STRMU
4	84	88.4	1592	1	GTFC_STRMU
5	73	76.8	1365	1	GTFC_STRMU
6	63	68.4	1462	1	GTFC_STRMU
7	43	45.3	214	1	CMPA_BACST
8	42	44.2	644	1	RNB_ECOLI
9	42	44.2	644	1	RNB_ECOLI
10	42	44.2	644	1	RNB_ECOLI
11	42	44.2	659	1	RNB_ECOLI
12	42	44.2	932	1	HRL_SCHPO
13	41.5	43.7	713	1	DCUZ_ECOLI
14	41	43.2	344	1	MIZD_BACST
15	41	43.2	644	1	RNB_SALTI
16	41	43.2	644	1	RNB_SALTI
17	41	43.2	658	1	RNB_PASMT
18	40	42.1	205	1	ADEN_ADEG8
19	40	42.1	431	1	CRTX_PANAN
20	40	42.1	436	1	THIC_CLOAB
21	40	42.1	548	1	LIP2_CANRU
22	40	42.1	549	1	LIP3_CANRU
23	40	42.1	549	1	LIP4_CANRU
24	40	42.1	572	1	PTUB_STAUV
25	40	42.1	6359	1	BACC_BACLI
26	39.5	41.6	192	1	CUPP_DROPS
27	39	41.1	403	1	COAT_BOOLV
28	39	41.1	509	1	CA43_DROME
29	39	41.1	557	1	WR20_ARATH
30	39	41.1	585	1	ASNS_WAIZE
31	39	41.1	715	1	DOA1_YEAST
32	39	41.1	737	1	SKN1_CANAL
33	39	41.1	946	1	YER6_YEAST

34	39	41.1	1383	1	PRAX_RAT	063425 rattus norv
35	39	41.1	2090	1	N214_HUMAN	P35658 homo sapien
36	39	41.1	2116	1	MY52_DICDI	P08799 dictyosteli
37	38.5	40.5	417	1	PREB_HUMAN	Q9HCU5 homo sapien
38	38.5	40.5	454	1	NEM_PIG	P08552 sus scrofa
39	38.5	40.5	499	1	AINX_HUMAN	Q16352 homo sapien
40	38.5	40.5	500	1	PTGI_HUMAN	Q16647 homo sapien
41	38.5	40.5	504	1	AINX_MOUSE	P46660 mus musculu
42	38.5	40.5	505	1	AINX_RAT	P23565 rattus norv
43	38.5	40.5	810	1	NEM_BOVIN	O77788 bos taurus
44	38.5	40.5	845	1	NEM_RAT	P12839 rattus norv
45	38.5	40.5	915	1	NEM_HUMAN	P07197 homo sapien

ALIGNMENTS

RESULT 1
GTFC_STRMU STANDARD: PRT: 1455 AA.
ID GTFC_STRMU
AC P13470; 069382; 069385; 069386; 069391; 069397; P05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
CN GTF OR SMU 1005
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RX STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terzo Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Pringleaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
[4]
RN SEQUENCE OF 1-349 FROM N.A.
RP STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=340685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 183:4263-4270(1987).
[5]
RN FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CARBOHYDRATE ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl- (N) = D-
fructose + (1,6)-alpha-D-glucosyl- (N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.

CC	-I-	SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.
CC	-I-	SIMILARITY: Contains 10 cell wall binding repeats.
CC	-- -- -- -- --	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; M17361;	AAA88588.1; ; -
DR	EMBL; D88651;	BAA26101.1; ; -
DR	EMBL; D88654;	BAA26105.1; ; -
DR	EMBL; D88657;	BAA26109.1; ; -
DR	EMBL; D88660;	BAA26113.1; ; -
DR	EMBL; D89977;	BAA26119.1; ; -
DR	EMBL; AE014940;	AAN58705.1; ; -
DR	InterPro; IPR002479;	CW binding.
DR	InterPro; IPR003318;	Glyco hydro_70.
DR	Pfam; PF01473;	CW binding_I; 1..1.
KM	pPfo2324;	Glyco hydro_70.1.
KM	transferase; Glycosyltransferase;	Signal; Repeat; Dental carrier;
KM	Complete proteome.	
FT	SIGNAL	1 34
FT	CHAIN	35 1476
FT	DOMAIN	35 1051
FT	DOMAIN	1097 1476
FT	REPEAT	1097 1130
FT	DOMAIN	1161 1470
FT	REPEAT	1161 1210
FT	REPEAT	1225 1275
FT	REPEAT	1290 1340
FT	REPEAT	1355 1405
FT	REPEAT	1420 1470
FT	VARIANT	62 62
FT	VARIANT	65 65
FT	VARIANT	68 68
FT	VARIANT	68 68
FT	VARIANT	78 78
FT	VARIANT	86 86
FT	VARIANT	89 89
FT	VARIANT	168 168
FT	VARIANT	276 276
FT	VARIANT	399 399
FT	VARIANT	474 474
FT	VARIANT	512 512
FT	VARIANT	519 519
FT	VARIANT	701 701
FT	VARIANT	708 708
FT	VARIANT	938 938
FT	VARIANT	952 957
FT	VARIANT	963 964
FT	VARIANT	968 970
FT	VARIANT	1086 1086
FT	VARIANT	1158 1158
FT	VARIANT	1163 1163
FT	VARIANT	1168 1168
FT	VARIANT	1182 1182
FT	VARIANT	1234 1234
FT	VARIANT	1263 1263
FT	VARIANT	1264 1264
FT	VARIANT	1272 1272
FT	VARIANT	1329 1329

```

FT  VARIANT 1394 1394 Y -> H (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT 1402 1402 S -> G (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT 1459 1459 Y -> H (IN STRAIN MT4467) .
FT  VARIANT 570 570 R -> A (IN REF. 1) .
FT  CONFLICT 800 817 ADDVRYVASTASTASTDCK -> LKMFALRLARPHQWA
FT  CONFLICT 1310 1310 (IN REF. 1) .
FT  CONFLICT 1476 AA; 165846 MM; 9CCE03F731B4CBFC CRC64;
SQ  SEQUENCE

Query Match 95.8%; Score 91; DB 1; Length 1476;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VPSSYFIRAHDSYVDLI 18
Db 552 VPSSYFIRAHDSYVDLI 569

RESULT 3
GTF1_STRDO STANDARD; PRT; 1597 AA.
ID_GTF1_STRDO
AC p11001;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteriia; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Gilpin M.L., Ruseell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFe28".
RL J. Bacteriol. 169:4271-4278 (1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS. THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1) .
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-S1 SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.stb.ch).
CC -----
CC EMBL; M17391; AAC63063.1; -
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CW binding_1; 16.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT 1 38 POTENTIAL.

```

```

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1597 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MM; B9B86A200868798B CRC64;

Query Match 90.5%; Score 86; DB 1; Length 1597;
Best Local Similarity 94.4%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPSYSTRANDSEVODLI 18
Db 554 VPSYSTRANDSEVODLI 571

RESULT 4
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohba H., Kato K.,
RA Kagawa H.;
RT "peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase)"; 173:989-996(1991).
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BAAL4241.1; -
CC DR InterPro; IPR002479; CW_binding.

```

```

DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT CHAIN 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MM; BC0A6D079351ECF CRC64;

Query Match 88.4%; Score 84; DB 1; Length 1592;
Best Local Similarity 88.9%; Pred. No. 5.5e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPSYSTRANDSEVODLI 18
Db 548 VPSYSTRANDSEVODLI 565

RESULT 5
GTF5_STRDO STANDARD; PRT; 1365 AA.
ID GTF5_STRDO
AC P29356;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTF5.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MFE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Kilmore K.S., Russell J.R., Ferretti J.J.;
RT "analysis of the Streptococcus downei gtfS gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans";
RL Infect. Immun. 58:2452-2458(1990).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30943; AAA26898.1; -
CC DR

```


DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF0324; Glyco_hydro_70; 1.
 KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 36
 FT CHAIN 37 1365 GLUCOSYLTRANSFERASE-S.
 FT REPEAT 157 177 CELL WALL BINDING 1.
 FT REPEAT 178 197 CELL WALL BINDING 2.
 FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
 FT REPEAT 1062 1082 CELL WALL BINDING 3.
 FT REPEAT 1083 1102 CELL WALL BINDING 4.
 FT REPEAT 1150 1169 CELL WALL BINDING 5.
 FT REPEAT 1170 1190 CELL WALL BINDING 6.
 FT REPEAT 1225 1243 CELL WALL BINDING 7.
 FT REPEAT 1289 1308 CELL WALL BINDING 8.
 FT REPEAT 1309 1328 CELL WALL BINDING 9.
 FT REPEAT 1331 1352 CELL WALL BINDING 10.
 SQ SEQUENCE 1365 AA; 151590 MW; 1672965A2B8C476 CRC64;
 Query Match 76.8%; Score 73; DB 1; Length 1365;
 Best Local Similarity 78.9%; Pred. No. 0.00034;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VPSYSEFIRAHDSVQDLIA 19
 Db 537 VPSYSEFIRAHDSVQDLIA 555
 RESULT 6
 GTFD_STRMU STANDARD; PRT; 1462 AA.
 ID GTFD_STRMU STANDARD; PRT; 1462 AA.
 AC P49331; O69383; O69386; O69389; O69392; O69398;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFD OR SMU.910.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=91100958; PubMed=2148600;
 RA Honda O., Kato C., Kuramitsu H.K.;
 RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
 the glucosyltransferase-S enzyme.";
 RL J. Gen. Microbiol. 136:2099-2105(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 MT4467 / Serotype E, and MT8148 / Serotype C;
 RX MEDLINE=98231643; PubMed=9570124;
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of
 Streptococcus mutans.";
 RL PEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 CC EMBL; M29296; AAA26895.1; -;
 CC EMBL; D88653; BAA26103.1; -;
 CC EMBL; D88656; BAA26107.1; -;
 CC EMBL; D88659; BAA26111.1; -;
 CC EMBL; D88662; BAA26115.1; -;
 CC EMBL; D89979; BAA26121.1; -;
 CC EMBL; AE014932; AAN58619.1; -;
 CC InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF0324; Glyco_hydro_70.1.
 KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 FT CHAIN 1 1462
 FT DOMAIN 1232 1423
 FT REPEAT 1232 1295
 FT REPEAT 1296 1359
 FT REPEAT 1360 1423
 FT VARIANT 10 10
 FT 19 19
 FT VARIANT 58 58
 FT VARIANT 68 68
 FT VARIANT 81 81
 FT VARIANT 113 113
 FT VARIANT 122 122
 FT 132 132
 FT VARIANT 135 135
 FT VARIANT 137 137
 FT 202 202
 FT VARIANT 255 255
 FT VARIANT 275 275
 FT 288 288
 FT 301 301
 FT VARIANT 313 313
 FT VARIANT 317 317
 FT VARIANT 328 328
 FT VARIANT 350 350
 FT 628 633
 FT VARIANT 688 688
 FT VARIANT 726 732
 FT VARIANT 726 730
 FT 762 762
 FT VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; BAA26103.1; -;
 EMBL; D88656; BAA26107.1; -;
 EMBL; D88659; BAA26111.1; -;
 EMBL; D88662; BAA26115.1; -;
 EMBL; D89979; BAA26121.1; -;
 EMBL; AE014932; AAN58619.1; -;
 InterPro; IPR002479; CW_binding.
 InterPro; IPR003318; Glyco_hydro_70.
 Pfam; PF0324; Glyco_hydro_70.1.
 Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 Complete proteome.
 CHAIN 1 1462
 DOMAIN 1232 1423
 REPEAT 1232 1295
 REPEAT 1296 1359
 REPEAT 1360 1423
 VARIANT 10 10
 19 19
 VARIANT 58 58
 VARIANT 68 68
 VARIANT 81 81
 VARIANT 113 113
 VARIANT 122 122
 132 132
 VARIANT 135 135
 VARIANT 137 137
 202 202
 VARIANT 255 255
 VARIANT 275 275
 288 288
 301 301
 VARIANT 313 313
 VARIANT 317 317
 VARIANT 328 328
 VARIANT 350 350
 628 633
 VARIANT 688 688
 VARIANT 726 732
 VARIANT 726 730
 762 762
 VARIANT 762 762
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 SUBCELLULAR LOCATION: Secreted.
 DISEASE: DENTAL CARIES.
 MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 SIMILARITY: Contains 6 cell wall binding repeats.

 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).

 EMBL; M29296; AAA26895.1; -;
 EMBL; D88653; B

```

FT VARIANT 964 964 D -> Y (IN STRAIN MT4251).
FT VARIANT 1019 1019 E -> K (IN STRAINS MT4245 AND MT4251).
FT VARIANT 1059 1059 LG -> IR (IN STRAIN MT4251).
FT VARIANT 1060 1060 G -> R (IN STRAIN MT4245).
FT VARIANT 1080 1080 G -> R (IN STRAIN MT4245).
FT VARIANT 1142 1142 H -> Q (IN STRAIN GS-5).
FT VARIANT 1198 1198 S -> N (IN STRAIN MT4239).
FT VARIANT 1220 1220 Y -> C (IN STRAINS MT4251 AND MT4467).
FT VARIANT 1280 1280 F -> L (IN STRAIN MT4467).
FT VARIANT 1282 1282 Q -> P (IN STRAIN MT4245).
FT VARIANT 1290 1290 K -> T (IN STRAIN MT4245).
FT VARIANT 1311 1311 N -> D (IN STRAIN MT4245).
FT VARIANT 1403 1403 G -> D (IN STRAIN GS-5 AND MT4467).
FT VARIANT 1425 1425 G -> R (IN STRAIN GS-5).
FT VARIANT 1449 1449 R -> K (IN STRAIN MT4467).
FT CONFLICT 1448 1449 RYDKNSGMVKNKVTLLNNGRRIGIDRWGIARY -> VY
FT CONFLICT 1462 1462 R (IN REF. 1).
SQ SEQUENCE 1462 AA; 163387 MW; CBA4279CAD708645 CRC64;

Query Match 68.4%; Score 65; DB 1; Length 1462;
Best Local Similarity 76.5%; Pred. No. 0.0085;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 19
Db 576 NYIFRAHDSVQTVIA 592

RESULT 7
CMPA_BACSU
ID ID PRT; 214 AA.
AC P14264; STANDARD;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Competence protein A.
GN COMA OR COMDA OR COMAT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90008771; PubMed=2507523;
RA Weinrauch Y., Guillen N., Dubnau D.;
RT "Sequence and transcription mapping of Bacillus subtilis competence
RT genes comB and comA, one of which is related to a family of bacterial
RT regulatory determinants."
RL J. Bacteriol. 171:5362-5375 (1989).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=97419515; PubMed=9274030;
RA Oudega B., Koningsstein G., Rodriguez L., de Sales Ramon M.,
RA Hilbert H., Duesterhoeft A., Pohl T.M., Weitzenecker T.;
RT "Analysis of the Bacillus subtilis genome: cloning and nucleotide
RT sequence of a 62 kb region between 275 degrees (rrnB) and 284 degrees
RT (gal).";
RL Microbiology 143:2769-2774 (1997).
[3]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Noszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Beterio M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortles R., Bourschi L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Comerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Duesterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabre C., Ferrari B., Foulger D.,
RA Fitt C., Fujita Y., Fuma S., Galizzi A., Galleon N.,
RA Gim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech U., Harwood C.R., Henaut A.,
RA Hilbert H., Holappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

```

```

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lander J., Lazarevic C.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigic C.,
RA Medina N., Mellado R.F., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Pardo V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
RA Priesack E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni K.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretti A.,
RA Viari A., Wandut R., Wedler E., Wedler H., Weitzenecker T.,
RA Winers P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256 (1997).
CC -!- FUNCTION: Required for the development of competence in
CC B.subtilis, for the expression of late-expressing competence
CC genes, and for the expression of the growth stage-regulated
CC molecule surfactin. May play a regulatory role during the
CC development of competence, and an analogous role as uvrC-ORF2
CC (transducing environmental information to the DNA repair system).
CC -!- PTM: Phosphorylated by comB.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -!- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL M22856; AAA22320.1; -
DR EMBL Z93932; CAB07904.1; -
DR EMBL Z99120; CAB15156.1; -
DR PIR A33591; RGSBCA.
DR Subtilist; BG10381; comA.
DR InterPro: IPR000792; HTH_LuxR.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00196; GcrE; 1.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00038; HTH_LuxR.
DR ProDom: PD000307; HTH_LuxR; 1.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00421; HTH_LuxR; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS00622; HTH_LuxR_FAMILY; 1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
DR Competence; Sensory transduction; Transcription regulation; Activator;
KM DNA-binding; Phosphorylation; Complete proteome.
KW DOMAIN 1 121 RESPONSE REGULATORY.
FT MOD_RES 55 55 PHOSPHORYLATION (BY SIMILARITY).
FT DNA_BIND 171 190 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 214 AA; 24128 MW; 52D01A9740759072 CRC64;

Query Match 45.3%; Score 43; DB 1; Length 214;
Best Local Similarity 52.9%; Pred. No. 5.4;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 PSYSFIRAHDSVQDLI 18
Db 36 PSEQFIKQHDPSYSYDLI 52

RESULT 8
RNB_ECO57

```


DR EMBL: D90766; BAA14840.1; -.
DR PIR: A64877; A64877.
DR EcoGene: EG11620; rnb.
DR HAMAP: MF_01036; -; 1.
DR InterPro: IPR004476; 3_prime_RNase.
DR InterPro: IPR002059; Cold_shock.
DR InterPro: IPR001900; Ribonuclease_II.
DR InterPro: IPR003029; SL.
DR Pfam: PF00773; RNB; 1.
DR Pfam: PF00575; SL; 1.
DR SMART: SM00357; CSP; 1.
DR SMART: SM00316; SL; 1.
DR TIGRfam: TIGR00358; 3_prime_RNase; 1.
DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
DR PROSITE: PS01126; SL; FALSE NEG.
DR HydroLase: Exonuclease; Nuclease; RNA-binding; Complete proteome.
DR DOMAIN 561 643
FT CONFLICT 384 384 I -> N (IN REF. 1).
FT CONFLICT 399 399 C -> G (IN REF. 1).
FT CONFLICT 513 513 A -> R (IN REF. 1).
FT SEQUENCE 644 AA; 72490 MW; 36B16712CDF14394 CRC64;

Query Match 44.2%; Score 42; DB 1; Length 644;
Best Local Similarity 44.4%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 PYSYFIRAHDSVQDLIA 19
DB 362 PDYRFLGKGEVLDIVA 379

RESULT 10
RNB_SHIFL ID STANDARD; PRT; 644 AA.
AC P59107;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II).
GN RNB OR SFI291.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
ON NCBI_TaxID=623;
RX STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Sun L., Xue Y., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Chang H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -1- FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded
CC polyribonucleotides processively in the 3' to 5' direction (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in the 3'- to 5'-
CC direction to yield nucleoside 5'-phosphates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ribonuclease II (RNB) family.
CC -1- SIMILARITY: Contains 1 SI motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE015155; AAN42902.1; -.
DR HAMAP: MF_01036; -; 1.
DR InterPro: IPR004476; 3_prime_RNase.
DR InterPro: IPR002059; Cold_shock.
DR InterPro: IPR001900; Ribonuclease_II.
DR InterPro: IPR003029; SL.
DR Pfam: PF00773; RNB; 1.
DR Pfam: PF00575; SL; 1.
DR SMART: SM00357; CSP; 1.
DR SMART: SM00316; SL; 1.
DR TIGRfam: TIGR00358; 3_prime_RNase; 1.
DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
DR PROSITE: PS01126; SL; FALSE NEG.
DR HydroLase: Exonuclease; Nuclease; RNA-binding.
DR DOMAIN 561 643
FT CONFLICT 384 384 I -> N (IN REF. 1).
FT CONFLICT 399 399 C -> G (IN REF. 1).
FT SEQUENCE 644 AA; 72464 MW; 137EAF724EF3735C2 CRC64;

Query Match 44.2%; Score 42; DB 1; Length 644;
Best Local Similarity 44.4%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 PYSYFIRAHDSVQDLIA 19
DB 362 PDYRFLGKGEVLDIVA 379

RESULT 11
RNB_HAEIN ID STANDARD; PRT; 659 AA.
AC P44440;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II).
GN RNB OR H11733.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
ON NCBI_TaxID=727;
RX STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton R., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,
RA Uettermann J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uettermann T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Grehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded
CC polyribonucleotides processively in the 3' to 5' direction (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in the 3'- to 5'-
CC direction to yield nucleoside 5'-phosphates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ribonuclease II (RNB) family.
CC -1- SIMILARITY: Contains 1 SI motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U32846; AAC23378.1; -.

```

DR PIR; A64139; A64139.
DR TIGR; H11733; -.
DR HAMAP; MF_01036; -. 1.
DR InterPro; IPR004476; 3_prime_RNase.
DR InterPro; IPR002059; Cold_shock.
DR InterPro; IPR001900; Ribonuclease_II.
DR InterPro; IPR003029; S1.
DR Pfam; PF00773; RNH; 1.
DR Pfam; PF00575; S1; 1.
DR SMART; SM00357; CSP; 1.
DR SMART; SM00316; S1; 1.
DR TIGRfams; TIGR00358; 3_prime_RNase; 1.
DR PROSITE; PS01175; RIBONUCLEASE_II; 1.
DR PROSITE; PS50126; S1; 1.
DR Hydrolase; Exonuclease; Nuclease; RNA-binding; Complete proteome.
KW DOMAIN 576 658 S1 MOTIF.
SQ SEQUENCE 659 AA; 75782 MW; FDDA9097A7D3B66 CRC64;

Query Match 44.2%; Score 42; DB 1; Length 659;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 PSYFIRAHSEVQDLIA 19
DB 367 PDYAFVLAENGKVOEIKK 384

RESULT 12
HTRL SCHPO STANDARD; PRT; 932 AA.
AC P87314;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone transcription regulator 1 homolog.
GN SPBC31F10.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymonopoulos B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Meller-auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Kestel D., Hilbert H.,
RA Borzym K., Zimmermann W., Wedler H., Leinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu F., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forstburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpukovskii G.V., Ussery D., Barrett B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe";
RT Nature 415:871-889(2002).
CC -!- FUNCTION: COULD PLAY A PART IN MECHANISMS OF TRANSCRIPTIONAL
REGULATION SIMILAR TO THAT PLAYED BY YEAST HIR1 AND HIR2 TOGETHER.

```

```

CC -!- SIMILARITY: Contains 8 WD repeats.
CC -!- SIMILARITY: BELONGS TO THE HIR1 FAMILY OF WD-REPEAT PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z97204; CAB10089.1; -.
CC PIR; T40216; T40216.
CC GeneDB; Spombe; SPBC31F10.13C; -.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 6.
CC PRINTS; PR00320; GPROTEINRPT.
CC ProDom; PD000018; WD40; 2.
CC SMART; SM00320; WD40; 6.
CC PROSITE; PS00678; WD_REPEATS_1; 1.
CC PROSITE; PS50082; WD_REPEATS_2; 4.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Transcription regulation; Repeat; WD repeat;
KW Nuclear protein.
FT REPEAT 16 55 WD 1.
FT REPEAT 72 111 WD 2.
FT REPEAT 132 171 WD 3.
FT REPEAT 174 213 WD 4.
FT REPEAT 268 316 WD 5.
FT REPEAT 320 361 WD 6.
FT REPEAT 623 663 WD 7.
FT REPEAT 664 709 WD 8.
SQ SEQUENCE 932 AA; 103686 MW; 09D527CDB9003D3E CRC64;

Query Match 44.2%; Score 42; DB 1; Length 932;
Best Local Similarity 46.7%; Pred. No. 40;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDL 17
DB 126 SYRRLGSHNDIQDL 140

RESULT 13
DCLZ ECOLI STANDARD; PRT; 713 AA.
AC P52035; P78299;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysine decarboxylase, constitutive (BC 4.1.1.16) (LDC).
GN LDC OR LDC OR LDC OR B0186.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=K12 / W3110;
RC MEDLINE=97369816; PubMed=9226257;
RX Kirsch Y., Kojima H., Tanaka T., Takatsuka Y., Kamio Y.;
RT "Characterization of a second lysine decarboxylase isolated from
RT Escherichia coli.";
RT J. Bacteriol. 179:4486-4492(1997).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40 AND 415-456.
RC STRAIN=K12 / W3110;
RC MEDLINE=97480327; PubMed=9339543;
RA Yamamoto Y., Miwa Y., Miyoshi K., Furuyama J., Ohmori H.;
RT "The Escherichia coli ldc gene encodes another lysine decarboxylase,
RT probably a constitutive enzyme.";
RT Genes Genet. Syst. 72:167-172(1997).
RN [3]

```

RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Samped G., Mizobuchi K.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of the
 RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Pederspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Lashbary D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
 RA Davis R.W.;
 RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 555-713 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=98369317; PubMed=9723924;
 RA Pichoff S., Alibaud L., Guedat A., Castanie M.-P., Bouche J.-P.;
 RT "An *Escherichia coli* gene (yaeO) suppresses temperature-sensitive
 RT mutations in essential genes by modulating Rho-dependent
 RT transcription termination.";
 RL Mol. Microbiol. 29:859-869(1998).
 CC -1- FUNCTION: IDC IS CONSTITUTIVELY BUT WEAKLY EXPRESSED UNDER VARIOUS
 CC CONDITIONS. OPTIMUM ACTIVITY IS ACHIEVED BETWEEN PH 6.2 TO 8.0.
 CC -1- CATALYTIC ACTIVITY: L-lysine = cadaverine + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- SUBUNIT: HOMODECAMER.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF ORNITHINE, LYSINE, AND ARGININE
 CC DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D87518; BAA21656.1; -;
 DR EMBL: D49445; BAA08426.1; -;
 DR EMBL: AEO00128; AAC73297.1; -;
 DR EMBL: D83536; BAA77861.1; -;
 DR EMBL: U70214; AAB08615.1; -;
 DR EMBL: Z50870; CAA90749.1; -;
 DR PIR: B64743; B64743.
 DR HSSP: P43099; IC4K.
 DR EcoGene: EG31219; IdCC.
 DR InterPro: IPR000310; Decarboxylase1.
 DR InterPro: IPR005108; OKR_DC_1_N.
 DR Pfam: PF01276; OKR_DC_1_1.
 DR Pfam: PF03711; OKR_DC_1_C_1.
 DR Pfam: PF03709; OKR_DC_1_N_1.
 DR PROSITE: PS00703; OKR_DC_1_1.
 KW Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
 KW Complete proteome.
 FT BINDING 367
 FT CONFLICT 284
 FT CONFLICT 314
 FT CONFLICT 411
 FT CONFLICT 413
 FT CONFLICT 498
 M -> T (IN REF. 2).
 N -> F (IN REF. 2).
 T -> S (IN REF. 2).
 AA -> R (IN REF. 2).
 M -> I (IN REF. 2).

FT CONFLICT 673 673 L -> I (IN REF. 2).
 SQ SEQUENCE 713 AA; 80590 MW; 4532C6069744BDF CRC64;
 Query Match 43.7%; Score 41.5; DB 1; Length 713;
 Best Local Similarity 45.0%; Pred. No. 36;
 Matches 9; Conservative 3; Mismatches 5; Indels 3; Gaps 1;
 QY 1 VPSYIFIRAH--DSVQDL 17
 DB 77 LPVAFINTHSTMDVSVDPM 96
 RESULT 14
 MIDZ BACSU STANDARD; PRT; 344 AA.
 ID MIDZ BACSU
 AC P26935;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 GN Myo-inositol 2-dehydrogenase (EC 1.1.1.18).
 GN IDH OR IOLG OR E83G.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / 60015;
 RX MEDLINE=92104493; PubMed=1761221;
 RA Fujita Y., Shindo K., Miwa Y., Yoshida K.;
 RT "Bacillus subtilis inositol dehydrogenase-encoding gene (idh):
 RT sequence and expression in *Escherichia coli*.";
 RL Gene 108:121-125(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSCL1;
 RX MEDLINE=95039891; PubMed=7952181;
 RA Yoshida K., Sano H., Miwa Y., Ogasawara N., Fujita Y.;
 RT "Cloning and nucleotide sequencing of a 15 kb region of the *Bacillus*
 RT subtilis genome containing the *iol* operon.";
 RL Microbiology 140:2289-2298(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunze F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertoletto M.G., Bessieres P., Bolotin A., Borcher S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano N.J., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertien K.D., Errington J., Fabret C., Ferrari B., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G., Hentut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulle M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Meinel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Nodack M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudga B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan B., Schleich S., Schroeter R., Scifone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vamler P., Vaasarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yara K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*

```

RT subtlis.";
RL Nature 390:249-256(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=79239346; PubMed=112095;
RA Ramealey R., Fujita Y., Freese E.;
RT "Purification and properties of Bacillus subtilis inositol
  dehydrogenase."
RL J. Biol. Chem. 254:7684-7690(1979).
CC -1- CATALYTIC ACTIVITY: Myo-inositol + NAD(+) = 2,4,6/3,5-
  pentahydroxycyclohexanone + NADH.
CC -1- PATHWAY: Myo-inositol catabolism; first step.
CC -1- SUBUNIT: Monomer.
CC -1- INDUCTION: INOSITOL. SUBJECTED TO CATABOLITE REPRESSION.
CC -1- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M76431; AAA2543.1; -.
DR EMBL: D14399; BAA03296.1; -.
DR EMBL: Z99124; CAB16006.1; -.
DR PIR: JH0511; JH0511.
DR SUBLIT: BG10669; 1dh.
DR InterPro: IPR000683; GFO_IDH_MOCA.
DR InterPro: IPR004104; GFO_IDH_MOCA_C.
DR Pfam: PF01408; GFO_IDH_MOCA; 1.
DR Pfam: PF02894; GFO_IDH_MOCA_C; 1.
DR Oxidoreductase; NAD; Complete proteome.
SQ SEQUENCE 344 AA; 38351 MW; 2FC890BD4E2C332F CRC64;

Query Match 43.2%; Score 41; DB 1; Length 344;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 FIRAHSEVQDLI 18
   |||:|:|:|
Db 276 FVAAYDVEIQDFI 288

RESULT 15
RNB_SALTI STANDARD; PRT; 644 AA.
ID RNB_SALTI
AC Q827C9;
DT 26-FEB-2003 (Rel. 41, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II).
GN RNB OR STY1350 OR T1615.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OC NCB1_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parikh J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
  Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
  Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
  Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
  Felzwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
  Krogsh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
  Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
  Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar typhi CT18.";
RL Nature 413:848-852(2001).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
  Burland V., Kodyanski V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar typhi strains Ty2
  and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded
  polyribonucleotides processively in the 3' to 5' direction (By
  similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in the 3'- to 5'-
  direction to yield nucleoside 5'-phosphates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ribonuclease II (RNB) family.
CC -1- SIMILARITY: Contains 1 SI motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL627270; CAD01619.1; -.
DR EMBL: AE016839; AAO69242.1; -.
DR HAMAP: MF_01036; -; 1.
DR InterPro: IPR004476; 3_prime_RNase.
DR InterPro: IPR002059; Cold shock.
DR InterPro: IPR001900; Ribonuclease_II.
DR InterPro: IPR003029; SI.
DR Pfam: PF00773; RNB; 1.
DR Pfam: PF00575; SI; 1.
DR SMART: SM00357; CSP; 1.
DR SMART: SM00316; SI; 1.
DR TIGRPFAM: TIGR00358; 3_prime_RNase; 1.
DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
DR PROSITE: PS50126; SI; FALSE NEG.
KW Hydrolyase; Exonuclease; Nuclease; RNA-binding; Complete proteome.
FT DOMAIN 561 643 SI MOTIF.
SQ SEQUENCE 644 AA; 72481 MW; 17AA90ACC4C8DEB CRC64;

Query Match 43.2%; Score 41; DB 1; Length 644;
Best Local Similarity 38.9%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 PSYSTRANDSEVQDLIA 19
   |||:|:|
Db 362 PDYRFLGKGEVLDIYA 379

Search completed: November 13, 2003, 09:45:27
Job time: 7.8436 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 28.6351 Seconds

(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049A-2

Perfect score: 95

Sequence: 1 VPSSYFIRAHDESEVQDLIA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	84	88.4	1590	2	Q59983 streptococc
2	84	88.4	1590	2	Q55263 streptococc
3	80	84.2	1477	2	Q91466 leuconostoc
4	80	84.2	1508	2	Q95285 leuconostoc
5	80	84.2	1508	2	Q52224 leuconostoc
6	80	84.2	1527	2	Q92AR4 leuconostoc
7	80	84.2	1527	2	Q8KRE1 leuconostoc
8	80	84.2	1554	2	Q8KZ15 streptococc
9	79	83.2	1290	2	Q48756 leuconostoc
10	77	81.1	1016	2	Q91CJ7 leuconostoc
11	75	78.9	2835	2	Q8G9Q2 streptococc
12	74	77.9	1512	2	Q9WXJ5 streptococc
13	73	76.8	1338	2	Q9WXJ4 streptococc
14	66	69.5	2057	2	Q9RE05 leuconostoc
15	65	68.4	1449	2	Q68542 streptococc
16	65	68.4	1449	2	Q55264 streptococc

17	65	68.4	1518	2	Q00600 streptococc
18	64	67.4	1575	2	Q9LCH3 streptococc
19	64	67.4	1577	2	Q54178 streptococc
20	64	67.4	1599	2	Q00599 streptococc
21	61	64.2	1577	2	Q55265 streptococc
22	50	52.6	93	2	Q92IX9 streptococc
23	47	49.5	480	5	Q95WU6 bacillus lam
24	46	48.4	66	2	Q9K5K6 bacillus mo
25	46	48.4	558	5	Q9GYN4 caenorhabdi
26	46	48.4	575	5	P90900 caenorhabdi
27	46	48.4	631	16	Q98P18 mycoplasma
28	45	48.4	1032	10	Q9F1J9 arabidopsis
29	46	48.4	1035	10	Q8G571 arabidopsis
30	45.5	47.9	248	16	Q987H1 rhizobium l
31	44	46.3	249	17	Q978J3 thermoplas
32	44	46.3	327	10	Q9C7I1 arabidopsis
33	44	46.3	371	8	Q9B8A8 trichinella
34	43	45.3	51	16	Q9RNP7 vibrio chol
35	43	45.3	64	2	Q9K5I0 bacillus su
36	43	45.3	590	17	Q9P5R0 methanobac
37	43	45.3	1260	10	Q9LMD9 arabidopsis
38	42	44.2	247	10	Q9M8T9 arabidopsis
39	42	44.2	431	2	Q47843 pantoea agg
40	42	44.2	466	12	Q8OLN2 arabidopsis
41	42	44.2	512	12	Q8JMF0 arabidopsis
42	42	44.2	563	16	Q9JZL7 neisseria m
43	42	44.2	563	16	Q9JUP8 neisseria m
44	42	44.2	644	16	Q8FHT6 escherichia
45	42	44.2	881	16	Q8G5W2 bifidobacte

ALIGNMENTS

ID	Q59983	PRELIMINARY	PRT: 1590 AA.
AC	Q59983		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)		
DE	Glucosyltransferase-I precursor (EC 2.4.1.5).		
GN	GTRF.		
OS	Streptococcus sobrinus.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
CC	Streptococcus.		
OX	NCBI_TaxID=1310;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=OM2176;		
RX	MEDLINE=94146405; PubMed=8312602;		
RA	Sato S., Inoue M., Handa N., Aizawa Y., Isebe Y., Katayama T.;		
RT	"DNA sequence of the glucosyltransferase gene of serotype d		
RT	Streptococcus sobrinus."		
RU	DNA Seq. 4:19-27(1993).		
DR	EMBL, D13858; BAO02976.1; -		
DR	InterPro; IPR002479; CW_binding.		
DR	InterPro; IPR00318; Glyco_hydro_70.		
DR	Pfam; PF02324; Glyco_hydro_70; 1.		
KW	Glycosyltransferase; Signal; Transferase.		
FT	SIGNAL 1 38		
FT	CHAIN 39 1590		
SQ	SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;		
QY	Query Match	88.4%; Score 84; DB 2; Length 1590;	
DB	Best Local Similarity	88.9%; Pred. No. 1.7e-05;	
	Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
	1 VPSSYFIRAHDESEVQDLI 18		
	548 VPSSYFIRAHDESEVQDLI 565		


```
RESULT 2
Q55263 ID Q9E263 PRELIMINARY; PRT; 1590 AA.
AC O55263;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE GTF-1.
GN GLUCOSYLTRANSFERASE.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase
produced from Streptococcus sobrinus ATCC 33478."
RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL; D63570; BAA09792.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 15.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6EAFD43 CRC64;

Query Match 88.4%; Score 84; DB 2; Length 1590;
Best Local Similarity 88.9%; Pred. No. 1.7e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDII 18
Db 548 VPSYSFIRAHSEVQDII 565

RESULT 3
Q9L466 ID Q9L466 PRELIMINARY; PRT; 1477 AA.
AC Q9L466;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Remaud-Simeon M., Piznut S., Sarcabal P.,
Willemt R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
RT glucosyltransferase from leuconostoc mesenteroides NRRL B-1355."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250112; CAB76565.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
KM GLYCOSYLTRANSFERASE; Transferase.
SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFCB831 CRC64;

Query Match 84.2%; Score 80; DB 2; Length 1477;
Best Local Similarity 73.7%; Pred. No. 8e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDII 19
Db 603 IPNYSFVRAHSEVQTVIA 621
```

```
RESULT 4
Q9E2H5 ID Q9E2H5 PRELIMINARY; PRT; 1508 AA.
AC Q9E2H5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA087AE4F3A CRC64;

Query Match 84.2%; Score 80; DB 2; Length 1508;
Best Local Similarity 73.7%; Pred. No. 8.1e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDII 19
Db 634 IPNYSFVRAHSEVQTVIA 652
```

```
RESULT 5
O52224 ID O52224 PRELIMINARY; PRT; 1508 AA.
AC O52224;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Remaud-Simeon M., Monsan P., Willemt R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan."
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
KM GLYCOSYLTRANSFERASE; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 84.2%; Score 80; DB 2; Length 1508;
Best Local Similarity 73.7%; Pred. No. 8.1e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDII 19
Db 634 IPNYSFVRAHSEVQTVIA 652

RESULT 6
```

```

Q9ZAR4          PRELIMINARY;      PRT; 1527 AA.
ID Q9ZAR4;
AC Q9ZAR4;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE Dextranucrase.
GN DEX.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512-F;
RA Bhattacharya R., Singh D.K.S.;
RT "Cloning and Molecular Characterization of Dextranucrase Gene from
RT Leuconostoc mesenteroides NRRL B-512F."
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; 081374; AAD10952.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1527 AA; 169709 MW; 1DFAFA237C743398 CRC64;

Query Match      84.2%; Score 80; DB 2; Length 1527;
Best Local Similarity 73.7%; Pred. No. 8.3e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIA 19
Db 652 IPNYSFVRADHSEVQTVIA 670

RESULT 7
Q8KRE1          PRELIMINARY;      PRT; 1527 AA.
ID Q8KRE1;
AC Q8KRE1;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Dextranucrase Dsrd (EC 2.4.1.5).
GN DSRD.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RA Neubauer H., Bauche A., Mollet B.;
RT "Isolation and characterization of the dextranucrase Dsrd of
RT Leuconostoc mesenteroides Lcc4."
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY017384; AAG61158.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glycosyltransferase.
KM TRANSFERASE; Glycosyltransferase.
SQ SEQUENCE 1527 AA; 169835 MW; F9D0DB220BD9668 CRC64;

Query Match      84.2%; Score 80; DB 2; Length 1527;
Best Local Similarity 73.7%; Pred. No. 8.3e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIA 19
Db 652 IPNYSFVRADHSEVQTVIA 670

RESULT 8
Q8KZL5          PRELIMINARY;      PRT; 1554 AA.
ID Q8KZL5;
AC Q8KZL5;

```

```

DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Glucosyltransferase.
GN GTFU.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21956884; PubMed=11960691;
RA Hanada N., Fukushima K., Nomura Y., Sengoku H., Hayakawa M.,
RA Mukasa H., Shirota T., Abiko Y.;
RT "Cloning and nucleotide sequence analysis of the Streptococcus
RT sobrinus gltF gene that produces a highly branched water-soluble
RT glucan."
RT Biochim. Biophys. Acta 1570:75-79(2002).
RL EMBL; AB089438; BAC07265.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW TRANSFERASE.
SQ SEQUENCE 1554 AA; 171676 MW; 6981BC1DAE24A73 CRC64;

Query Match      84.2%; Score 80; DB 2; Length 1554;
Best Local Similarity 73.7%; Pred. No. 8.4e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIA 19
Db 557 IPTYSFVRADHSEVQTVIA 575

RESULT 9
Q48756          PRELIMINARY;      PRT; 1290 AA.
ID Q48756;
AC Q48756;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B1299;
RX MEDLINE=97136686; PubMed=8982063;
RA Monchois V., Willemot R.M., Renaud-Simeon M., Croux C., Monsan P.;
RT "Cloning and sequencing of a gene coding for a novel dextranucrase
RT from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-
RT 6) and alpha (1-3) linkages."
RL Gene 182:23-32(1996).
RU EMBL; U38181; AAB40875.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 11.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1290 AA; 145590 MW; 3555C2E96B749FAA CRC64;

Query Match      83.2%; Score 79; DB 2; Length 1290;
Best Local Similarity 83.3%; Pred. No. 0.0001;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PYSYFIRAHDSVQDLIA 19
Db 388 PNYSFIRAHDSVQTVIA 405

RESULT 10
Q9LCT7

```


DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)
DE Alernansucrase (EC 2.4.1.140).
GN ASR.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacilliales; Leuconostoc.
OX NCBI_TaxId=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RX MEDLINE=20080809; PubMed=10612736;
RA Aquello-Morales M.A., Remaud-Simeon M., Pizunt S., Sarahabi P.,
RA Willenot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alernansucrase, a sucrose
RT glycosyltransferase from leuconostoc mesenteroides NRRL B-1355.";
RL FEMS Microbiol. Lett. 182:81-85(2000).
DR EMBL; AJ250173; CAB65910.2; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003118; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 12.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 2057 AA; 228987 MW; 62BCE9385D9A11BE CRC64;

Query Match	69.5%	Score 66	DB 2	Length 2057
Best Local Similarity	61.1%	Pred. No.	0.031	
Matches 11, Conservative	4	Mismatches	3	Indels 0; Gaps 0

```
QY      1 VPSYSFIRAHDSVQDLI 18
        :|:|:|:|:|:|:|:|:|
Db      757 IPNYSFVRADYDAQDFI 774
```

```

RESULT 15
068542
AC 068542; PREDIMINARY; PRT; 1449 AA.
DT 01-NOV-1998 (rEMBLrel. 07, Created)
DT 01-NOV-1998 (rEMBLrel. 07, last sequence update)
DT 01-OCT-2002 (rEMBLrel. 22, last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases
DR EMBL, AF049609, AAC05156.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
FT NON TER. 1449 1449
SQ SEQUENCE 1449 AA; 159695 MW; 0700FED748471BF8 CRC64

```

Query Match	68.4%	Score 65	DB 2	Length 1449
Best Local Similarity	68.8%	Score No. 0.032		
Matches 11	Conservative 4	Mismatches 1	Indels 0	Gaps 0

```
QY      3 SYSFIRAHDSVQDLI 18
        :|:|:|:|:|:|:|:|
Db      609 NYAFVRAHDSEVQSTI 624
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 40.3507 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049a-3

Sequence: 1 TGAATGQLLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: A_Geneseq_19Jun03:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	23	AAU98027
2	112	100.0	1475	23	AAU98030
3	112	100.0	1475	23	AAU98031
4	112	100.0	1475	23	AAU98032
5	112	100.0	1475	23	AAU98033
6	112	100.0	1475	23	AAU98034
7	112	100.0	1475	23	AAU98035
8	112	100.0	1475	23	AAU98036
9	112	100.0	1475	23	AAU98037

10	112	100.0	1475	23	AAU98038	S. mutans GTFB mut
11	112	100.0	1475	23	AAU98039	S. mutans glucosyl
12	112	100.0	1475	23	AAU98040	S. mutans glucosyl
13	105	93.8	1476	23	AAU98041	Streptococcus muta
14	105	93.8	1476	23	AAU98042	Streptococcus muta
15	87	77.7	22	14	AAU98043	GTF subsequence. S
16	87	77.7	22	18	AAU98044	GTF antigenic pep
17	87	77.7	22	18	AAU98045	GTF antigenic epit
18	87	77.7	24	18	AAU98046	GTF antigenic pep
19	87	77.7	26	14	AAU98047	Carboxyltransfera
20	87	77.7	1592	14	AAU98048	Glucosyltransfera
21	81	72.3	1375	23	AAU98049	Dextran saccharase
22	78	69.6	1375	23	AAU98050	S. mutans glucosyl
23	78	69.6	1375	23	AAU98051	S. mutans glucosyl
24	67	59.8	1430	23	AAU98052	S. mutans glucosyl
25	67	59.8	1430	23	AAU98053	S. mutans glucosyl
26	67	59.8	1430	23	AAU98054	S. mutans glucosyl
27	67	59.8	1430	23	AAU98055	S. mutans glucosyl
28	67	59.8	1430	23	AAU98056	S. mutans glucosyl
29	67	59.8	1430	23	AAU98057	S. mutans glucosyl
30	66	58.9	1577	17	AAU98058	Alpha-D-glucosyltr
31	65	58.0	2057	21	AAU98059	L. mesenteroides a
32	57	50.9	1527	23	AAU98060	Leuconostoc mesent
33	54.5	48.7	32	23	AAU98061	Dextran-saccharas
34	49.5	44.2	320	20	AAU98062	Toxin A immunogeni
35	49.5	44.2	457	20	AAU98063	Toxin A immunogeni
36	49.5	44.2	811	17	AAU98064	C. difficile toxin
37	49.5	44.2	812	17	AAU98065	C. difficile toxin
38	49.5	44.2	862	21	AAU98066	Toxin A immunogeni
39	49.5	44.2	866	21	AAU98067	A recombinant prot
40	49.5	44.2	866	21	AAU98068	C. difficile toxin
41	49.5	44.2	2710	19	AAU98069	Clostridium difficile
42	49.5	44.2	2710	19	AAU98070	P. falciparum yf8p
43	49	43.8	343	23	AAU98071	P. falciparum isopr
44	49	43.8	343	23	AAU98072	Human polypeptide
45	47	42.0	765	22	AAU98073	

ALIGNMENTS

```

RESULT 1
ID AAU98027 standard; Protein; 1475 AA.
XX
AC AAU98027;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB.
XX
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture.
XX
OS Streptococcus mutans.
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX

```


SQ Sequence 1475 AA;
Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGARTINGQLLYFRANGVQVKG 22
1300 TGARTINGQLLYFRANGVQVKG 1321
Db 1300 TGARTINGQLLYFRANGVQVKG 1321
RESULT 3
AAU98031
ID AUA98031 standard; Protein; 1475 AA.
XX
XX AAU98031;
XX
XX 27-AUG-2002 (first entry)
XX
XX S. mutans glucosyltransferase GTFB mutant D457N.
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 457 /note="Wild-type Asp substituted by Asn"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX WPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substitutes for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
XX (GTF) B polypeptide having changes at position from 1448V, D457N,
XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
XX D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790C/K1014T,
XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
XX changes at positions from T589P, T589E, N471D, N471D/T589D, and
XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
XX an isolated polynucleotide which encodes P1 or P2, or its complementary
XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
XX an expression cassette comprising the polynucleotide operably linked to a
XX promoter, a vector comprising the expression cassette, host cell
XX introduced with the vector, a transgenic plant comprising the
XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
XX coating composition comprising a glucan produced in a plant transformed
XX with a gene encoding the mutant GTF, wild type or, starch, a latex,

thermoplastic molecule or their combinations or glucan and starch where
the glucan is produced in the amyloplast and/or vacuole or a maize line
deficient in starch biosynthesis, transformed with a gene encoding a
glucosyltransferase B or D enzyme, wild-type or mutant and a paper
comprising the glucan (paper sizing/coating agent). The vector is useful
for producing a glucan in a plant. The method comprises transforming a
plant cell with the vector, growing the plant cell under plant growing
conditions to produce a regenerated plant and inducing expression of the
polynucleotide for a time sufficient to produce the glucan in the
regenerated plant, where the vector contains a transit sequence from
ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
chlorophyll AB binding protein to produce a transgenic plant, and glucan
is produced in the amyloplast of potato or the vacuole of sugar beet.
Glucans are useful as substitutes for and additions to modified starch
and latexes in paper manufacture. Unlike prior art techniques, which
require input materials that produce chemical effluents, paper
manufacture utilizing the glucan produced by GTF, which utilizes
biologically produced input materials, is more cost-effective and
environmentally friendly. Moreover, glucans also exhibit thermoplastic
properties and impart gloss to the paper during coating step.
The present sequence represents a GTFB mutant of the invention.
Note: The present sequence is not shown in the specification but
was created by the indexer using the GTFB sequence appearing as AAU98027
and the information in claim 36.
SQ Sequence 1475 AA;
Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGARTINGQLLYFRANGVQVKG 22
1300 TGARTINGQLLYFRANGVQVKG 1321
Db 1300 TGARTINGQLLYFRANGVQVKG 1321
RESULT 4
AAU98032
ID AUA98032 standard; Protein; 1475 AA.
XX
XX AAU98032;
XX
XX 27-AUG-2002 (first entry)
XX
XX S. mutans glucosyltransferase GTFB mutant D567N.
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 567 /note="Wild-type Asp substituted by Thr"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX

PI Nicholas;
XX
XX
XX WPI, 2002-41332/44.
XX
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substrates for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -: 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF). B. polypeptide having changes at position from 1445H, D457N,
CC D567T, K1014I, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1445H/D457N/D567T/D571K/K793Q/K1014T,
CC D169A/Y170W/Y171A, and K479Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589E, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to
CC a promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll a/b binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and/or additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as ANU98027
CC and the information in claim 36.

SQ	Sequence	1475 AA;
Query Match	100.0%;	Score 112; DB 23; Length 1475;
Best Local Similarity	100.0%;	Pred. No. 3.8e-10;
Matches 22;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0

[illegible]

RESULT 5	
AAU98033	
ID	AAU98033 standard; Protein; 1475 AA.

AC	AAU98033;
XX	
DT	27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant K1014T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing; coating composition; glucan; starch; latex; thermoplastic molecular

KW	amy1op1ase; vacuole; paper manufacture; mutant; mutein.
XX	
OS	Streptococcus mutans.
OS	Synthetic.
XX	
XX	Key
FH	Misc-difference
FT	1014
XX	/note= "Wild-type Lys substituted by Thr"
PN	US2002031826-A1.

PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274

PR 11-DEC-1998: 98US-0210361
PR 07-JUN-1993: 95US-0478704
PR 07-JUN-1995: 95US-0482711
PR 07-JUN-1993: 95US-0485243
PR 16-JAN-1998: 98US-0007599
PR 16-JAN-1998: 98US-0008172
PR 20-JAN-1998: 98US-0009620

PA (NICH/) NICHOLS S E

PI Nichols SE;

DR WPT; 2002-414332/44

PT Glucosyltransferase

PT paper manufacture, c

PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
DA

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having

N471D/T589E. Also included are a glucan produced by the GTF

polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising one or more genes, and a promoter, a vector comprising the expression cassette, host cell

vector, a seed or tuber from the transgenic plant, a paper sizing and/or

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC coating composition comprising a glucan produced in a plant

the glycan is produced in the amyloplast and/or vacuole or a maize line

deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme wild-type or mutant and a paper

comprising the glucan (paper sizing/coating agent): the vector is used for producing a clone. The method comprises transforming a

plant cell with the vector, growing the plant cell under plant growing

polynucleotide for a time sufficient to produce the glucan in the

CC ribulose biphosphate carboxylase small subunit, waxy, D111E-1 and

CC is produced in the amyloplast of potato or the vacuole of sugar beec.

CC and latexes in paper manufacture. Unlike prior art techniques, which

CC manufacture utilising the glucan produced by GTF, which utilises

environmentally friendly. Moreover, glucans also exhibit thermoplastic

The present sequence represents a GTFB mutant of the invention.

CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.
XX Sequence 1475 AA;
SQ
Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGARTINGOLLYPRANGVQVKG 22
Db 1300 TGARTINGOLLYPRANGVQVKG 1321
RESULT 6
AAU98034 ID AAU98034 standard; Protein: 1475 AA.
XX AC AAU98034;
XX 27-AUG-2002 (first entry)
DT
XX S. mutans glucosyltransferase GTFB mutant D457N/D567T.
DB
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutain.
XX Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT
XX US2002031826-A1.
XX 14-MAR-2002.
XX 19-DEC-2000; 2000US-0740274.
XX 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
PI Nichols SE;
XX
DR WPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
PS Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF B polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes PI or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex.
CC thermoplastic molecule or their combinations of glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
XX
Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGARTINGOLLYPRANGVQVKG 22
Db 1300 TGARTINGOLLYPRANGVQVKG 1321
RESULT 7
AAU98035 ID AAU98035 standard; Protein: 1475 AA.
XX AC AAU98035;
XX 27-AUG-2002 (first entry)
DT
XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.
DB
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutain.
XX Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT
XX US2002031826-A1.
XX 14-MAR-2002.
XX 19-DEC-2000; 2000US-0740274.
XX 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a plant transformed
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 XX
 QY Query Match 100.0%; Score 112; DB 23; Length 1475;
 DB Best Local Similarity 100.0%; Pred. No. 3 8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1300 TGARTINGQLLYFRANGVQYKG 1321
 RESULT 8
 AAU98036
 ID AAU98036 standard; Protein; 1475 AA.
 XX
 AC AAU98036;
 XX

DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567 /note= "wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "wild-type Asp substituted by Thr"
 FT FT
 FT
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impact gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TGARTINGQLLYFRANGVQVKG 22
 Db 1300 TGARTINGQLLYFRANGVQVKG 1321
 RESULT 9
 ID AAU98037 standard; Protein: 1475 AA.
 AC AAU98037;
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impact gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TGARTINGQLLYFRANGVQVKG 22
 Db 1300 TGARTINGQLLYFRANGVQVKG 1321
 RESULT 10
 ID AAU98038 standard; Protein: 1475 AA.
 AC AAU98038;
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567

FT /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779
 FT /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014
 FT /note= "Wild-type Lys substituted by Thr"
 FT US2002031826-A1.
 PD 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT Claim 36; Page -; 44pp; English.
 PS
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, D11E1-1 and
 CC chloroplast AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer, using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX SQ Sequence 1475 AA;
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQVKG 22
 DB 1300 TGARTINGQLLYFRANGVQVKG 1321
 RESULT 11
 ID AAU98039
 AAU98039 standard; Protein; 1475 AA.
 AC AAU98039;
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 169..171
 FT /note= "Wild-type Tyr-Tyr-Tyr substituted by
 FT Ala-Ala-Ala"
 XX
 XX US2002031826-A1.
 XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT Claim 36; Page -; 44pp; English.
 PS
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGARTINGQLLYFRANGVQVKG 22
1300 TGARTINGQLLYFRANGVQVKG 1321

RESULT 12

AAU98040
ID AAU98040 standard; Protein; 1475 AA.

AAU98040;

27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant K779Q.

Glucosyltransferase; GTFB; transgenic plant; paper sizing; coating composition; glucan; starch; latex; thermoplastic molecule; amyloplast; vacuole; paper manufacture; mutant; mutain.

Streptococcus mutans.

Synthetic.

Key Location/Qualifiers

Misc-difference 779 /note= "Wild-type Lys substituted by Gln"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

11-DEC-1998; 98US-0210361.

07-JUN-1995; 95US-0478704.

07-JUN-1995; 95US-0482711.

16-JAN-1998; 98US-0007939.

20-JAN-1998; 98US-0009620.

(NICH/) NICHOLS S. E.

Nichols SE;

WPI: 2002-414332/44.

Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions - Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T, Y159A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGARTINGQLLYFRANGVQVKG 22
1300 TGARTINGQLLYFRANGVQVKG 1321

RESULT 13

AAU79286
ID AAU79286 standard; Protein; 380 AA.

AAU79286;

13-AUG-2002 (first entry)

Streptococcus mutans monoclonal antibody-related protein #3.

KW Antibody; dental caries; water insoluble glucan synthetase;
 XX anti-carries; glucosyl transferase-B; immunotherapy.
 XX Streptococcus mutans.
 OS JP2002114709-A.
 XX PD 16-APR-2002.
 XX PF 04-OCT-2000; 2000JP-0304889.
 XX PR 04-OCT-2000; 2000JP-0304889.
 XX PA (UYNI-) UNIV NIPPON.
 XX DR WPI; 2002-448101/48.
 PT Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX PS Claim 5; Page 19-20; 28pp; Japanese.
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC SQ Sequence 380 AA;
 Query Match 93.8%; Score 105; DB 23; Length 380;
 Best Local Similarity 95.5%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TGARTINGQLTYFRANGVQVKG 22
 DB 205 TGARTINGQLTYFRANGVQVKG 226
 RESULT 14
 AAU79284
 ID AAU79284 standard; Protein; 1476 AA.
 XX AAU79284;
 AC 13-AUG-2002 (first entry)
 DT Streptococcus mutans monoclonal antibody-related protein #1.
 XX Streptococcus mutans monoclonal antibody-related protein #1.
 XX Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-carries; glucosyl transferase-B; immunotherapy.
 XX Streptococcus mutans.
 OS JP2002114709-A.
 XX PD 16-APR-2002.
 XX PF 04-OCT-2000; 2000JP-0304889.
 XX PR 04-OCT-2000; 2000JP-0304889.
 XX PA (UYNI-) UNIV NIPPON.
 XX DR WPI; 2002-448101/48.
 PT Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX PS Claim 3; Page 13-16; 28pp; Japanese.
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC SQ Sequence 1476 AA;
 Query Match 93.8%; Score 105; DB 23; Length 1476;
 Best Local Similarity 95.5%; Pred. No. 6.3e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TGARTINGQLTYFRANGVQVKG 22
 DB 1301 TGARTINGQLTYFRANGVQVKG 1322
 RESULT 15
 AAAR43696
 ID AAAR43696 standard; peptide; 22 AA.
 XX AAAR43696;
 AC 25-MAR-2003 (updated)
 DT 20-MAY-1994 (first entry)
 XX GT subsequence.
 DE GT; glucosyltransferase; vaccine; T-cell; B-cell; reaction;
 KW immunoreponse; peptidyl core matrix; dental caries; diptheria;
 KW tetanus; measles; polio.
 XX Synthetic.
 OS WO9322341-A1.
 XX WO9322341-A1.
 XX 11-NOV-1993.
 PD 30-APR-1993; 93WO-US04094.
 XX 01-MAY-1992; 92US-0877295.
 XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
 XX Smith DJ, Taubman MA;
 XX WPI; 1993-368721/46.
 DR Synthetic glucosyl-transferase peptide(s) - provoke T- and B-cell
 XX reactions in mammals, and are effective against dental caries
 XX PT Claim 3; Page 23; 38pp; English.
 XX The sequences (AAAR43694-98) are subsequences from GT, they provoke
 CC T- and B-cell responses in mammals. The can be used to produce a
 CC vaccine comprising 2 of these sequences attached to a peptidyl
 CC core matrix. They are useful in producing T-cell responses to
 CC the GT enzyme causing a reduction of colonisation or accumulation
 CC of mutans streptococcal strains in mammals. The vaccines can be used
 CC in preventing dental caries.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC SQ Sequence 22 AA;

Query Match 77.7%; Score 87; DB 14; Length 22;
 Best Local Similarity 77.3%; Pred. No. 4.6e-08;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGARTINGOLLYPRANGOVKG 22
 |||:|||||
 Db 1 TGAQTIKGQKLYFKANGQOVKG 22

Search completed: November 13, 2003, 09:38:26
 Job time : 41.3507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40, Search time 13.6588 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-3
Perfect score: 112
Sequence: 1 TGAATNGQLLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 76:*

1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	2 B3135	glbF protein precu
2	87	77.7	1592	2 A38175	glucosyltransferas
3	82	73.2	1599	2 S22737	glucosyltransferas
4	80	71.4	1518	2 A44811	glucosyltransferas
5	78	68.6	1375	2 UT0345	dextranucrase /BC
6	75	67.0	1290	2 UC5473	dextranucrase /BC
7	75	67.0	1365	2 A41483	glucosyltransferas
8	73	65.2	1449	2 T30857	glucosyltransferas
9	72	64.3	1508	2 T31098	probable dextran
10	68	60.7	563	2 A37184	glucan-binding pro
11	68	60.7	1449	2 T30552	glucosyltransferas
12	67	59.8	1431	2 A45866	dextranucrase /BC
13	66	58.9	1577	2 T30858	glucosyltransferas
14	65.5	58.5	329	2 A55221	dextranase inhibito
15	55	49.1	2817	2 B97033	uncharacterized pr
16	54	48.2	221	2 T37244	GATP transcrip:ion
17	50	44.6	831	2 T00323	chitinase (EC 3.2.
18	49.5	44.2	529	2 T10388	hypothetical prote
19	49.5	44.2	2710	2 A37052	tox A - Clostrid
20	49	43.8	491	2 C97069	cobyrilic acid synth
21	47.5	42.4	240	2 F97067	dethiobiotin synth
22	47	42.0	126	2 S40016	doc protein - phag
23	47	42.0	696	2 A28635	transcription fact
24	47	42.0	788	2 US0747	regulatory protein
25	46	41.1	231	2 F70471	conserved hypotet
26	46	41.1	245	2 F64465	hypothetical prote
27	46	41.1	301	2 S13165	asialoglycoprotein
28	46	41.1	619	2 A41971	surface protein ps
29	46	41.1	619	2 A97887	surface protein ps

30	46	41.1	744	2 F95013	pneumococcal surfa
31	46	41.1	1301	2 S18118	alpha-amylase - Al
32	46	41.1	2364	2 T40884	cyclooxin V - Clo
33	45	40.2	142	2 F87265	conserved hypotet
34	45	40.2	630	2 T05433	hypothetical prote
35	45	40.2	648	2 S10869	enterotoxin A - Cl
36	44	39.3	158	2 A12117	hypothetical prote
37	44	39.3	329	2 F71014	probable dipeptide
38	44	39.3	522	2 D71074	probable methylmal
39	44	39.3	522	2 H75135	methylmalonyl-coa
40	44	39.3	566	2 S74633	high affinity sulf
41	44	39.3	591	2 C84220	propionyl-CoA carb
42	44	39.3	601	2 AG0066	probable AMP-bindi
43	44	39.3	829	2 T29372	hypothetical prote
44	44	39.3	950	2 B87611	TonB-dependent rec
45	44	39.3	1396	2 S36851	L-shaped tail fibre

ALIGNMENTS

RESULT 1
B3135
glbF protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C:Accession: B3135; A33128
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the glbF gene from Streptococcus mutans.
A:Reference number: A33135; MUID:87308013; PMID:3040685
A:Accession: B3135
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1475 <SH1>
A:Cross-references: CB:M17361; NID:G153639; PIDN:AAA8588.1; PID:G153640
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
Submitted to the Protein Sequence Database, September 1990
A:Reference number: A33128
A:Accession: A33128
A:Status: Preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-171,173-641,'N',643-1475 <SH2>
A:Experimental source: strain GS-5
C:Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 100.0%; Score 112; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAATNGQLLYFRANGVQVKG 22
DB 1300 TGAATNGQLLYFRANGVQVKG 1321

RESULT 2
A38175
glucosyltransferase precursor - Streptococcus sobrinus
C:Species: Streptococcus sobrinus
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
C:Accession: A38175
R:Abu, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
J. Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A:Reference number: A38175; MUID:91123227; PMID:1704006
A:Accession: A38175
A:Status: Preliminary
A:Molecule type: DNA

A/Residues: 1-1592 <ABO>
 A/Cross-references: GB:D90213; NID:G217032; PIDN:BAAL241.1; PID:dl014946; PID:G217033
 C/Superfamily: cpl repeat homology
 F:1093-1112/Domain: cpl repeat homology <CP1>
 F:1222-1241/Domain: cpl repeat homology <CP2>
 F:1287-1306/Domain: cpl repeat homology <CP3>
 F:1330-1351/Domain: cpl repeat homology <CP4>
 F:1352-1371/Domain: cpl repeat homology <CP5>
 F:1402-1420/Domain: cpl repeat homology <CP6>
 F:1465-1484/Domain: cpl repeat homology <CP7>
 F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 77.7%; Score 87; DB 2; Length 1592;
 Best Local Similarity 77.3%; Pred. No. 8.9e-06;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1298 TGAQITRGQKLYFKANGQYKG 1319

RESULT 3
 S22737
 glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
 C/Accession: S22737; S28810; B44811; S22727
 R/Jacques, N.
 Submitted to the EMBL Data Library, March 1992
 A/Reference number: S22726
 A/Accession: S22737
 A/Molecule type: DNA
 A/Residues: 1-1599 <JAC>
 A/Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAA77698.1; PID:G47531
 A/Experimental source: ATCC 25975
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
 A/Reference number: A44811; MUID:92148377; PMID:1838391
 A/Accession: S28810
 A/Molecule type: DNA
 A/Residues: 1-51 <GLF>
 A/Cross-references: EMBL:Z11873
 C/Genetics: gtfK
 A/Superfamily: cpl repeat homology
 C/Keywords: glucosyltransferase; hexosyltransferase
 F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 73.2%; Score 82; DB 2; Length 1599;
 Best Local Similarity 77.3%; Pred. No. 6e-05;
 Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1402 TGAQVINGQLLYFDANGVQYKG 1423

RESULT 4
 A44811
 glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
 C/Accession: A44811; S22726; S28809
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase ge
 A/Reference number: A44811; MUID:92148377; PMID:1838391
 A/Accession: A44811
 A/Molecule type: DNA
 A/Residues: 1-1518 <GLF>
 A/Cross-references: EMBL:Z11873; NID:G47526; PIDN:CAA77900.1; PID:G47527
 A/Note: sequence extracted from NCBI backbone (NCBIN:81050, NCSTP:81052)

C/Genetics: gtfJ
 A/Gene: gtfJ
 C/Superfamily: cpl repeat homology
 C/Keywords: glycosyltransferase; hexosyltransferase
 F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 71.4%; Score 80; DB 2; Length 1518;
 Best Local Similarity 72.7%; Pred. No. 0.00012;
 Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1383 TGAQVINGQLLYFDANGVQYKG 1404

RESULT 5
 J0345
 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
 N/Alternate names: sucrose 6-glucosyltransferase
 C/Species: Streptococcus mutans
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
 C/Accession: J0345; C33135
 R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
 Gene 69, 101-109, 1988
 A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
 A/Reference number: J0345; MUID:89137980; PMID:2976010
 A/Accession: J0345
 A/Molecule type: DNA
 A/Residues: 1-1375 <UED>
 A/Experimental source: GS-5
 R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A/Reference number: A33135; MUID:87308013; PMID:3040685
 A/Accession: C33135
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-349 <SH1>
 A/Cross-references: GB:M17361
 C/Genetics: gtfC
 A/Gene: gtfC
 C/Punctuation:
 A/Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 C/Superfamily: cpl repeat homology
 C/Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:135-1375/Product: glucosyltransferase #status predicted <MAT>
 F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1253-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 69.6%; Score 78; DB 2; Length 1375;
 Best Local Similarity 68.2%; Pred. No. 0.00023;
 Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1264 TGTVTENGQRLYFKPVGQYKG 1285

RESULT 6
 J03473
 dextranucrase (EC 2.4.1.5) - Leuconostoc mesenteroides
 C/Species: Leuconostoc mesenteroides
 C/Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C/Accession: J03473
 R/Monchois, V.; Millemot, R.M.; Remaud-Simeon, M.; Croux, C.; Monsan, P.
 Gene 182, 23-32, 1996
 A/Title: Cloning and sequencing of a gene coding for a novel dextranucrase from Leuconoc
 A/Reference number: J03473; MUID:9713686; PMID:8982063
 A/Accession: J03473
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA

A;Residues: 1-1290 <MON>
A;Cross-references: GB:U38181
C;Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose onto C;Genetics:
A;Gene: dera
C;Keywords: glycosyltransferase; hexosyltransferase
F;78-870/Domain: catalytic #status predicted <CAT>
F;922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 67.0%; Score 75; DB 2; Length 1290;
Best Local Similarity 68.2%; Pred. No. 0.00068;
Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1245 TGVQVINGQLLYFRANGVQVKG 1266

RESULT 7

A41483
glucosyltransferase (EC 2.4.1.-) gtf5 precursor - Streptococcus sobrinus
C;Species: Streptococcus sobrinus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C;Accession: A41483
R;Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.U.
Infected. Immun. 58, 2452-2458, 1990
A;Title: Analysis of the Streptococcus downei gtf5 gene, which specifies a glucosyltrans
A;Reference number: A41483; MUID:90316665; PMID:2142479
A;Accession: A41483
A;Molecule type: DNA
A;Residues: 1-1365 <GIL>
A;Cross-references: GB:M30943; NID:g153652; PIDN:AAA26898.1; PID:g153653
C;Genetics:
A;Gene: gtf5
C;Superfamily: cpl repeat homology
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 67.0%; Score 75; DB 2; Length 1365;
Best Local Similarity 63.6%; Pred. No. 0.00072;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1235 TGEQTIDGQKVFQDNGVQVKG 1256

RESULT 8

T30857
glucosyltransferase - Streptococcus salivarius
C;Species: Streptococcus salivarius
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30857
R;Stimpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infected. Immun. 63, 609-621, 1995
A;Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A;Reference number: Z20909; MUID:95122197; PMID:7822030
A;Accession: T30857
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: DNA
A;Residues: 1-1449 <SIM>
A;Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAC41412.1
C;Genetics:
A;Gene: gtfL

Query Match 65.2%; Score 73; DB 2; Length 1449;
Best Local Similarity 68.2%; Pred. No. 0.0016;
Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1258 TGHQNINGQLLYFRANGVQVKG 1279

RESULT 9
T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides

C;Species: Leuconostoc mesenteroides
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C;Accession: T31098
R;Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A;Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A;Reference number: Z20981; MUID:98164374; PMID:9503626
A;Accession: T31098
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: DNA
A;Residues: 1-1508 <MON>
A;Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AAB95453.1
A;Experimental source: strain NRRL B-1239
C;Genetics:
A;Gene: dera
C;Function:
A;Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 64.3%; Score 72; DB 2; Length 1508;
Best Local Similarity 68.2%; Pred. No. 0.0025;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1399 TGLVNINGQLLYFRANGVQVKG 1420

RESULT 10

A37184
glucan-binding protein - Streptococcus mutans
C;Species: Streptococcus mutans
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 15-Oct-1999
C;Accession: A37184
R;Banat, J.A.; Russell, R.R.B.; Ferretti, J.U.
Infected. Immun. 58, 667-673, 1990
A;Title: Sequence analysis of the gene for the glucan-binding protein of Streptococcus m
A;Reference number: A37184; MUID:90170123; PMID:2307516
A;Accession: A37184
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-563 <BAN>
A;Cross-references: GB:M30945; NID:g153637; PIDN:AAA26894.1; PID:g153638
C;Superfamily: cpl repeat homology
F;169-188/Domain: cpl repeat homology <CP1>
F;264-283/Domain: cpl repeat homology <CP2>
F;349-368/Domain: cpl repeat homology <CP3>
F;504-523/Domain: cpl repeat homology <CP4>
F;525-548/Domain: cpl repeat homology <CP5>

Query Match 60.7%; Score 68; DB 2; Length 563;
Best Local Similarity 66.7%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GARTINGQLLYFRANGVQVKG 22
DB 276 GWRITGKXYFQDNGVQVKG 296

RESULT 11

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C;Species: Streptococcus salivarius
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30552
R;Jaffe, R.I.
submitted to the EMBL Data Library, February 1998
A;Description: Streptococcus salivarius V1477 gtfN.
A;Reference number: Z20854
A;Accession: T30552

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <JAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C:Genetics:
A:Gene: gtfN

Query Match 60.7%; Score 68; DB 2; Length 1449;
Best Local Similarity 66.7%; Pred. No. 0.011;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 GARTINGOLLYFRANGVQVKG 22
DB 1259 GHQNINGDELFEDNNQVQVKG 1279

RESULT 12

A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans

C:Species: Streptococcus mutans
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A45866

R:Honda, O.; Kato, C.; Kunitatsu, H.K.

J:Gen. Microbiol. 136, 2099-2105, 1990

A:Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl

A:Reference number: A45866; MUID:91100958; PMID:2148600

A:Accession: A45866

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1431 <HON>

A:Cross-references: GB:M29296

C:Superfamily: cpl repeat homology

C:Keywords: glycosyltransferase; hexosyltransferase

F:181-201/Domain: cpl repeat homology <CP1>

F:1127-1146/Domain: cpl repeat homology <CP2>

F:1192-1211/Domain: cpl repeat homology <CP3>

F:1257-1276/Domain: cpl repeat homology <CP4>

F:1277-1297/Domain: cpl repeat homology <CP5>

F:1321-1340/Domain: cpl repeat homology <CP6>

F:1341-1361/Domain: cpl repeat homology <CP6>

F:1385-1404/Domain: cpl repeat homology <CP7>

QY 1 TGARTINGOLLYFRANGVQVKG 22

DB 1332 TGSQTLAGKLYFASDGKQVKG 1353

Query Match 59.8%; Score 67; DB 2; Length 1431;
Best Local Similarity 59.1%; Pred. No. 0.016;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

RESULT 13

T30858
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858

R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1577 <SIM>

A:Cross-references: EMBL:J35928; NID:g662380; PID:g662381; PIDN:AAC41413.1

C:Genetics:

A:Gene: gtfM

Query Match 58.9%; Score 66; DB 2; Length 1577;
Best Local Similarity 54.5%; Pred. No. 0.026;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQVKG 22
DB 1469 TGLQNNKRYTFGSNGAQVKG 1490

RESULT 14

A55221

dextranase inhibitor precursor - Streptococcus sobrinus (strain UAB66, serotype g)

C:Species: Streptococcus sobrinus

C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 17-Mar-1999

C:Accession: A55221

R:Sun, J.W.; Wanda, S.Y.; Camilli, A.; Curtiss III, R.

J: Bacteriol. 176, 7213-7222, 1994

A:Title: Cloning and DNA sequencing of the dextranase inhibitor gene (dei) from Streptoc

A:Reference number: A55221; MUID:95050304; PMID:7961493

A:Accession: A55221

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-329 <STUN>

A:Cross-references: GB:L34406

C:Superfamily: cpl repeat homology

F:173-192/Domain: cpl repeat homology <CP1>

Query Match 58.5%; Score 65.5; DB 2; Length 329;
Best Local Similarity 68.2%; Pred. No. 0.0056;
Matches 15; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 GARTINGOLLYFRAN-GVQVKG 22
DB 185 GLQTINGQTLFEDNTGQVKG 206

RESULT 15

B97033
uncharacterized protein, related to enterotoxins of other Clostridiales [imported] - Clo

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: B97033

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: B97033

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2817 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79053.1; PID:G15023993; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1079

Query Match 49.1%; Score 55; DB 2; Length 2817;
Best Local Similarity 45.5%; Pred. No. 3.2;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQVKG 22
DB 1469 TGLQNTDNTTYFDSNGIMQVKG 1490

Search completed: November 13, 2003, 09:50:24
Job time: 15.6588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.32417 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-3

Perfect score: 112
Sequence: 1 TGARTINGQLLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	93.8	1476	1	GTFB_STRMU
2	87	77.7	1592	1	GTF2_STRDO
3	87	77.7	1597	1	GTF1_STRDO
4	86	76.8	1455	1	GTF3_STRMU
5	75	67.0	1365	1	GTF5_STRDO
6	67	59.8	1462	1	GTFD_STRMU
7	49.5	44.2	529	1	Y119_NPYOP
8	49.5	44.2	2710	1	TOXA_CLODI
9	47.5	42.4	240	1	BIOD_CLOAB
10	47	42.0	126	1	DOC_BPP1
11	47	42.0	781	1	SP1_MOUSE
12	47	42.0	785	1	SP1_MOUSE
13	47	42.0	788	1	SP1_MOUSE
14	46	41.1	231	1	DAH_AQUAE
15	46	41.1	231	1	USP_AQUAE
16	46	41.1	245	1	TM4_METUA
17	46	41.1	301	1	LCBI_MOUSE
18	45	40.2	587	1	MCCB_ARATH
19	44	39.3	1396	1	VITF_BPTS
20	43	38.4	110	1	YC27_ARCVU
21	43	38.4	467	1	YQ02_CAEEL
22	43	38.4	467	1	ISBP_SCHPO
23	41.5	37.1	401	1	HAT2_YEAST
24	41.5	37.1	472	1	CHIR_PELCA
25	41.5	37.1	472	1	CHIR_HUMAN
26	41.5	37.1	473	1	CHIR_MOUSE
27	41.5	37.1	473	1	CHIR_MOUSE
28	41.5	37.1	567	1	DY13_CHURE
29	41.5	37.1	579	1	COB2_BRARE
30	41	36.6	336	1	VINT_BP186
31	41	36.6	412	1	DP42_BACSU
32	41	36.6	412	1	DP41_BACSD
33	41	36.6	433	1	MPRC_PORGI

34	41	36.6	473	1	CBIR_TARGR	Q9PU17 taricha gra
35	41	36.6	513	1	Y073_TREPA	O83112 treponema p
36	41	36.6	924	1	OSTA_PSEAE	Q91502 pseudomonas
37	40.5	36.2	464	1	VL2_HPV27	P36755 human papil
38	40.5	36.2	524	1	VL2_HPV27	P25487 human papil
39	40	35.7	205	1	3MGE_CLOPE	O8349 clostridium
40	40	35.7	394	1	NUCC_STVY3	P27724 synochocyst
41	40	35.7	460	1	ALN_YEAST	P23375 saccharomyc
42	40	35.7	467	1	RFAL_CRIFA	Q23696 crithidia t
43	40	35.7	691	1	TOP1_BACSU	P39814 bacillus su
44	40	35.7	719	1	WZC_SALTY	O82566 salmonella
45	40	35.7	719	1	WZC_SALTY	Q9F7B1 salmonella

ALIGNMENTS

RESULT 1
GTFB_STRMU STANDARD; PRT; 1476 AA.
ID GTFB_STRMU
AC P08987; O69381; O69384; O69387; O69390; O69396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 15-SEP-2003 (Rel. 42, last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=67308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfb gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Jia S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Perreletti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
RC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
RC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
RC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N+1) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

```

CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17361; AAA86568.1; -
DR EMBL; D88651; BAA26101.1; -
DR EMBL; D88654; BAA26105.1; -
DR EMBL; D88657; BAA26109.1; -
DR EMBL; D88660; BAA26113.1; -
DR EMBL; D89977; BAA26119.1; -
DR EMBL; AE014940; AAN58705.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro.70.
DR Pfam; PF01473; CW_binding.1; 1.
DR Pfam; PF02324; Glyco_hydro.70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW Complete proteome.
KW
FT SIGNAL 1 34
FT CHAIN 35 1476
FT DOMAIN 35 1051
FT DOMAIN 1097 1476
FT REPEAT 1097 1130
FT DOMAIN 1161 1470
FT REPEAT 1161 1210
FT REPEAT 1225 1275
FT REPEAT 1290 1340
FT REPEAT 1355 1405
FT REPEAT 1420 1470
FT REPEAT 62 62
FT REPEAT 65 65
FT REPEAT 68 68
FT VARIANT 78 78
FT VARIANT 86 86
FT VARIANT 89 89
FT VARIANT 169 169
FT VARIANT 276 276
FT VARIANT 399 399
FT VARIANT 474 474
FT VARIANT 512 512
FT VARIANT 519 519
FT VARIANT 701 701
FT VARIANT 708 708
FT VARIANT 938 938
FT VARIANT 952 957
FT VARIANT 963 964
FT VARIANT 968 970
FT VARIANT 1086 1086
FT VARIANT 1158 1158
FT VARIANT 1163 1163
FT VARIANT 1168 1168
FT VARIANT 1182 1182
FT VARIANT 1234 1234
FT VARIANT 1263 1263
FT VARIANT 1263 1263
FT VARIANT 1264 1264
FT VARIANT 1272 1272
FT VARIANT 1329 1329
FT VARIANT 1394 1394
FT VARIANT 1402 1402

```

FT	VARIANT	1459	1459		Y -> H (IN STRAIN MT4467) .
FT	CONFLICT	570	570		R -> A (IN REF. 1) .
FT	CONFLICT	800	817		ADDDVRAATATASTDGK -> LIRKFALRLARPHQOMA
FT	CONFLICT				(IN REF. 1) .
FT	SEQUENCE	1310	1310		H -> I (IN REF. 1) .
SQ	SEQUENCE	1476 AA;	165846 MW;	9C6E09F731B4CBFC	CRC64;
Query Match			93.8%;	Score 105;	DB 1; Length 1476;
Best Local Similarity			95.5%;	Pred. No. 7,7e-09;	
Matches	21;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
Oy		1	TGARTINGOLLIYPFRANGVOYKG	22	
Dn		1301	TGARTINGOHLYPFRANGVOYKG	1322	
RESULT 2					
GTF2_STRDO					
ID	GTF2_STRDO	STANDARD;	PRT;	1592 AA.	
AC	P27470;				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)				
DS	(Sucrose 6-glucosyltransferase) .				
OS	Streptococcus downei (Streptococcus sobrinus) .				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
CC	Streptococcus.				
OX	NCHI_taxID=1317;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=6715 / Serotype G,				
RX	MEDLINE=911123227; PubMed=1704006;				
RA	Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,				
RA	Kagawa H.;				
RT	"Peptide sequences for sucrose splitting and glucan binding within				
RT	Streptococcus sobrinus glucosyltransferase (water-insoluble glucan				
RT	synthetase)." ;				
RL	J. Bacteriol. 173:989-996(1991) .				
CC	- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT				
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE				
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE				
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.				
CC	- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-				
CC	fructose + {(1,6)-alpha-D-glucosyl}(N+1) .				
CC	- SUBCELLULAR LOCATION: Secreted.				
CC	- DISEASE: DENTAL CARIES.				
CC	- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA				
CC	1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES				
CC	WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH				
CC	FORMS OF GLUCANS.				
CC	- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-				
CC	- BINDING PROTEIN FROM S.MUTANS.				
CC	- SIMILARITY: Contains 16 cell wall binding repeats.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on it's way				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-stb.ch/announce/				
CC	or send an email to license@isb-stb.ch).				
CC	-----				
DR	EMBL; D90213; BAAl4241.1; -				
DR	InterPro; IPR002479; CW_binding.				
DR	InterPro; IPR003318; Glyco_hydro_70.				
DR	Pfam; PF01473; CW_binding_I; 13.				
KM	Pfam; PF02324; Glyco_hydro_70; 1.				
FT	Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.				
FT	SIGNAL	1	38	POTENTIAL.	
FT	CHAIN	39	1592	GLUCOSYLTRANSFERASE-I.	
FT	DOMAIN	39	1044	CATALYTIC (APPROXIMATE) .	

```

FT DOMAIN 1093 1592          GLUCAN-BINDING (APPROXIMATE)
FT REPEAT 1093 1592          6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142          1.
FT REPEAT 1158 1207          2.
FT REPEAT 1222 1272          3.
FT REPEAT 1287 1337          4.
FT REPEAT 1402 1451          5.
FT REPEAT 1514 1563          6.
FT REPEAT 1577 1592          7. (INCOMPLETE)
SQ SEQUENCE 1592 AA; 176167 MM; BCOA66D079351ECF CRC64;

Query Match 77.7%; Score 87; DB 1; Length 1592;
Best Local Similarity 77.3%; Pred. No. 6,66-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGARTINGQILYFRANGVQVKG 22
Db 1298 TGAQITRGQKLYFRANGQGVVKG 1319

RESULT 3
ID GTF1_STRDO STANDARD; PRT, 1597 AA.
AC p11001;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (sucrose 6-glucosyltransferase).
GN GTF1.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEEZ8;
RA MEDLINE=87308014; PubMed=3040686;
RA Ferrerelli J.J., Gilpin M.L., Russell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MEEZ8";
RL J. Bacteriol. 169:4271-4278 (1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL, M17391; AAC63063.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.

```

```

FT FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-1.
FT FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT REPEAT 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1132 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 77.7%; Score 87; DB 1; Length 1597;
Best Local Similarity 77.3%; Pred. No. 6.6e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0

Db 1303 TGAQTIKGQKLYFRANGQVVG 1324
1 TGARTINGQLLYFRANGQVVG 22

RESULT 4
GTFc_STRMU STANDARD; PRT; 1455 AA.
AC P13470.069382; 069385; 069388; 069391; 069397; P05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN GTFc OR SMU.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 63:101-109(1988).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=MT439 / Serotype C; MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RX Fujitwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RN SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT

```

9T	VARIANT	1292	1294		DOH -> NOY (IN STRAINS GS-5, MT4467 AND MT8148).		
FT	VARIANT	1305	1369		MISSING (IN STRAIN MT4245).		
FT	VARIANT	1326	1326		I -> V (IN STRAIN MT8148).		
FT	VARIANT	1331	1331		T -> A (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).		
FT	VARIANT	1377	1377		R -> K (IN STRAIN MT8148).		
FT	VARIANT	1398	1398		V -> I (IN STRAIN MT8148).		
FT	VARIANT	1424	1424		D -> N (IN STRAIN MT4239).		
FT	VARIANT	1439	1439		V -> I (IN STRAINS MT4239 AND MT8148).		
FT	VARIANT	1444	1444		S -> P (IN STRAIN MT8148).		
FT	CONFLICT	1337	1455		ORLYPSNGVQAQGEILTERGKIKXYDDPNSGENEVNRYR TSGNMWYPENDCIVLGSIVREGRYRPDENGVYASD QRNHMDYDRDDGRSSSSAVPRHRANGFDFPF -> HASILSMFERLESSLOSXKVYSNTMILLPEAKFYI VM (IN REF. 1).		
9Q	SEQUENCE	1455 AA;	162965 MW;	3CB455A99A4FE0C86 CRC64;			
	Query Match		76.8%;	Score 86;	DB 1;		
	Best Local Similarity		68.2%;	Pred. No.	8.7e-06;		
	Matches	15;	Conservative	3;	Mismatches	4;	
				Indels	0;	Gaps	0;
Qy	1 TGARTINGOLLYFRANGOVYKG 22	: :					
Db	1329 TGTIRVANGQRITFKSNGVQAKG 1350	: :					
	RESULT 5						
GTF5_STRDO	ID	GTF5_STRDO	STANDARD;	PTI;	1365 AA.		
AC	P29336;						
DT	01-DEC-1992	(Rel. 24,	Created)				
DT	01-DEC-1992	(Rel. 24,	Last sequence update)				
DT	28-FEB-2003	(Rel. 4),	Last annotation update)				
DE	Glucosyltransferase-S precursor	(EC 2.4.1.5)	(GTF-S) (Dextranucrase)				
DE	(Sucrose 6-glucosyltransferase).						
GN	GTF5.						
OS	Streptococcus downei	(Streptococcus sobrinus).					
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;						
CC	Streptococcus;						
OX	NCBI_TaxID=1317;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=MFE28;						
RA	MEDLINE=90316665; PubMed=2142479;						
RT	Glucose K.S., Russell R.R., Ferretti J.J.;						
RT	"Analysis of the Streptococcus downei gtf5 gene, which specifies a						
RT	glucosyltransferase that synthesizes soluble glucans.";						
RL	Infect. Immun. 58:2452-2458 (1990).						
CC	- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT						
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE						
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE						
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.						
CC	- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-						
CC	fructose + {(1,6)-alpha-D-glucosyl} (N+1).						
CC	- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF						
CC	PRIMER GLUCAN UNLIKE GTF-1.						
CC	- DISASS: DENTAL CARIES.						
CC	- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA						
CC	1,6-GULOSE).						
CC	- SIMILIARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-						
CC	BINDING PROTEIN FROM S. MUTANS.						
CC	- SIMILIARITY: Contains 10 cell wall binding repeats.						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						
CC	use by non-profit institutions as long as its content is in no way						
CC	modified and this statement is not removed. Usage by and for commercial						
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/						
CC	or send an email to license@isb-sib.ch).						
CC	-----						
DR	EMBL; M30943; AAA26898.1; -.						

DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 36
 FT CHAIN 37 1365 OR 37 (POTENTIAL).
 FT REPEAT 157 177 GLUCOSYLTRANSFERASE-S.
 FT REPEAT 178 197 CELL WALL BINDING 1.
 FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
 FT REPEAT 1062 1082
 FT REPEAT 1083 1102 CELL WALL BINDING 3.
 FT REPEAT 1150 1169 CELL WALL BINDING 5.
 FT REPEAT 1170 1190 CELL WALL BINDING 6.
 FT REPEAT 1225 1243 CELL WALL BINDING 7.
 FT REPEAT 1289 1308 CELL WALL BINDING 8.
 FT REPEAT 1309 1328 CELL WALL BINDING 9.
 FT REPEAT 1331 1352 CELL WALL BINDING 10.
 FT SIGNAL 1365 AA; 151590 MW; 1672965A2E8C476 CRC64;
 SQ SEQUENCE

Query Match 67.0%; Score 75; DB 1; Length 1365;
 Best Local Similarity 63.6%; Pred. No. 0.00048;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQULYFRANGVQVG 22
 Db 1235 TGEOTIDGKVFQDNGVQVG 1256

RESULT 6
 GTFD_STRMU STANDARD; PRT; 1462 AA.
 AC P4331; O69383; O69386; O69389; O69392; O69398;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 26-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFD OR SMU. 910.
 OS Streptococcus mutans.
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=91100958; PubMed=2148600;
 RA Honda O., Kato C., Kuramitsu H.K.;
 RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
 the glucosyltransferase-S enzyme.";
 RL J. Gen. Microbiol. 136:2099-2105(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 RC MT4467 / Serotype E, and MT8148 / Serotype C;
 RX MEDLINE=98231643; PubMed=9570124;
 RA Fujiwara T., Teruo Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of
 Streptococcus mutans.";
 RL FEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdidi D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -!- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DISEASE: DENTAL CARIES.
 CC -!- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6-LINKAGES). GTF-S SYNTHESIZES BOTH
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -!- SIMILARITY: Contains 6 cell wall binding repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M29296; AAA26895.1; -;
 CC EMBL; D88653; BAA26103.1; -;
 CC EMBL; D88656; BAA26107.1; -;
 CC EMBL; D88658; BAA26111.1; -;
 CC EMBL; D88662; BAA26115.1; -;
 CC EMBL; D89979; BAA26121.1; -;
 CC EMBL; A8014932; AAN58619.1; -;
 CC InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 FT SIGNAL 1 1462
 FT CHAIN 1 1462
 FT DOMAIN 1232 1423
 FT REPEAT 1232 1295
 FT REPEAT 1296 1355
 FT REPEAT 1360 1423
 FT REPEAT 10 10
 FT VARIANT 19 19
 FT VARIANT 58 58
 FT VARIANT 68 68
 FT VARIANT 81 81
 FT VARIANT 113 113
 FT VARIANT 122 122
 FT VARIANT 132 132
 FT VARIANT 135 135
 FT VARIANT 137 137
 FT VARIANT 202 202
 FT VARIANT 255 255
 FT VARIANT 275 275
 FT VARIANT 288 288
 FT VARIANT 301 301
 FT VARIANT 313 313
 FT VARIANT 317 317
 FT VARIANT 328 328
 FT VARIANT 350 350
 FT VARIANT 628 633
 FT VARIANT 688 688
 FT VARIANT 726 732
 FT VARIANT 762 762
 FT VARIANT 762 762

POTENTIAL.
 GLUCOSYLTRANSFERASE-S.
 3 X 63 AA APPROXIMATE TANDEM REPEATS.
 1.
 2.
 3.
 Y -> H (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 I -> V (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 K -> E (IN STRAIN MT4467).
 A -> S (IN STRAINS MT4239 AND MT4245).
 A -> T (IN STRAINS MT4251 AND MT8148).
 T -> I (IN STRAINS MT4239 AND MT4245).
 A -> V (IN STRAINS MT4239, MT4245 AND
 MT8148).
 A -> S (IN STRAINS GS-5 AND MT4467).
 A -> V (IN STRAIN MT4245).
 A -> T (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 V -> L (IN STRAIN MT4239).
 D -> N (IN STRAIN MT8148).
 E -> D (IN STRAINS MT4239, MT4245 AND
 MT4251).
 D -> N (IN STRAINS MT4239, MT4245 AND
 MT4251).
 O -> H (IN STRAIN MT4245).
 D -> H (IN STRAINS MT4239 AND MT4251).
 E -> K (IN STRAIN MT4239).
 V -> F (IN STRAIN MT4239).
 F -> L (IN STRAINS MT4239, MT4251 AND
 MT4467).
 KKXYTO -> EKEYTL (IN STRAIN MT4251).
 A -> S (IN STRAIN MT4239).
 TDGSEA -> ADKNGS (IN STRAIN MT4251).
 TDGSS -> ADKGN (IN STRAINS MT4239 AND
 MT4245).
 T -> A (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).


```

FT  VARIANT  964  964  D -> Y (IN STRAIN MT4251).
FT  VARIANT  1019 1019  E -> K (IN STRAINS MT4245 AND MT4251).
FT  VARIANT  1059 1060  LG -> IR (IN STRAIN MT4251).
FT  VARIANT  1060 1060  G -> R (IN STRAIN MT4245).
FT  VARIANT  1080 1080  G -> R (IN STRAIN MT4239).
FT  VARIANT  1142 1142  H -> Q (IN STRAIN GS-5).
FT  VARIANT  1198 1198  S -> N (IN STRAIN MT4239).
FT  VARIANT  1220 1220  Y -> C (IN STRAINS MT4251 AND MT4467).
FT  VARIANT  1280 1280  Y -> L (IN STRAIN MT4467).
FT  VARIANT  1282 1282  Q -> P (IN STRAIN MT4245).
FT  VARIANT  1290 1290  K -> T (IN STRAIN MT4245).
FT  VARIANT  1311 1311  N -> D (IN STRAIN MT4245).
FT  VARIANT  1403 1403  G -> D (IN STRAINS GS-5 AND MT4467).
FT  VARIANT  1425 1425  G -> R (IN STRAIN GS-5).
FT  VARIANT  1449 1449  R -> K (IN STRAIN MT4467).
FT  CONFLICT 1428 1462  RYDKNSGNMVKVNLVLANGRGIDRMGIARY -> VY
SQ  SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;

Query Match 59.8%; Score 67; DB 1; Length 1462;
Best Local Similarity 59.1%; Pred. No. 0.01;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 1 TGARTINGOLLYFRANGOVYKG 22
Db 1332 TGSQTIAGKKLYFASDQKOVKG 1353

RESULT 7
Y19 NPVOP STANDARD; PRT; 529 AA.
ID Y19 NPVOP
AC 010358;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical 59.0 kDa protein precursor (ORF119).
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
[1]
RP MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohkman G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome."
RL Virology 229:381-399(1997).
CC -1- SIMILARITY: TO CORRESPONDING ORF IN ACNMPV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U75930; AAC5918.1; -.
CC DR Pfam; PF05092; DUF686; 1.
CC DR Hypothetical protein; Signal.
CC FT SIGNAL 1 20
CC FT CHAIN 21 529 HYPOTHETICAL PROTEIN ORF119.
CC FT SEQUENCE 529 AA; 58983 MW; E85F81DD9219BCB CRC64;

Query Match 44.2%; Score 49.5; DB 1; Length 529;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 11; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Oy 1 TGARTINGOLLYFRANGOVYKG 22
Db 260 SGRRT-SGRLEFYHADGVEVSG 280

```

```

RESULT 8
TOXA CLODI STANDARD; PRT; 2710 AA.
ID TOXA CLODI
AC P16154;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Toxin A.
GN TOXA OR TCDA.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
[1]
RP SEQUENCE FROM N.A.
RP STRAIN=VPI 10463;
RC MEDLINE=90221894; PubMed=2109310;
RX Sauerborn M., von Eichel-Streiber C.;
RA "Nucleotide sequence of Clostridium difficile toxin A."
RT Nucleic Acids Res. 18:1629-1630(1990).
[2]
RP SEQUENCE FROM N.A.
RP STRAIN=VPI 10463;
RC MEDLINE=90129305; PubMed=2105276;
RX Dore C.H., Wang S.Z., Price S.B., Phelps C.J., Lysterly D.M.,
RA Wilkins T.W., Johnson J.L.;
RT "Molecular characterization of the Clostridium difficile toxin A
RT gene."
RL Infect. Immun. 58:480-488(1990).
[3]
RP SEQUENCE FROM N.A.
RP STRAIN=VPI 10463;
RA von Eichel-Streiber C.;
RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- DOMAIN: THE C-TERMINAL PART OF TOXIN A CONSISTS OF A 83 AA
CC REPETITIVE STRUCTURE. THIS PART OF TOXIN A IS COMPOSED OF FIVE
CC DIFFERENT OLIGOPEPTIDES.
CC -1- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN
CC ENTEROTOXIN CALLED A AND CYTOTOXIN B. ONLY AFTER THE ENTERAL
CC DELIVERY OF THE ENTEROTOXIN A MAY THE CHARACTERISTIC DISEASE
CC CALLED PSEUDOMEMBRANOUS COLITIS BE INDUCED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X51797; CAA36094.1; -.
CC DR EMBL; M30307; AAA23283.1; -.
CC DR EMBL; X92982; CAA63564.1; -.
CC DR InterPro; IPR002479; CW binding.
CC DR Pfam; PF01473; CW binding 1; 28.
CC DR Pfam; PF04488; Gly_transf_sug; 1.
CC DR Toxin; Enterotoxin.
CC FT SEQUENCE 2710 AA; 308052 MW; 0A6E52CE84C14421 CRC64;

Query Match 44.2%; Score 49.5; DB 1; Length 2710;
Best Local Similarity 52.2%; Pred. No. 13;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Oy 1 TGARTIDGKKYFNTVAVATG 2462
Db 2460 TGARTIDGKKYFNTVAVATG 2462

RESULT 9
BIOD_CLOAB STANDARD; PRT; 240 AA.
ID BIOD_CLOAB
AC Q97UC5;

```

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dethiobiotin synthetase (EC 6.3.3.3) (Dethiobiotin synthase) (DTB
 GN synthetase) (DTBS).
 GN BIOT OR CAC1361.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1488;
 RX SEQUENCE FROM N.A.
 RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=1146286;
 RA Noelling J., Britton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gladson R.L., Sabathe F., Doucet-Stamm L., Soucaille P., Daly M.J.,
 RA Tatusov R.L., Koonin E.V., Smith D.R.,
 RA Bennett G.N., Koonin E.V., Smith D.R.,
 RA "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RT Bacteriol. 183:4823-4838(2001).
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -1- CATALYTIC ACTIVITY: ADP + 7,8-diaminononanoate + CO(2) = ADP +
 CC Phosphate + dethiobiotin.
 CC -1- COFACTOR: Magnesium (by similarity).
 CC -1- PATHWAY: Biotin conversion of pimelate into dethiobiotin.
 CC -1- SIMILARITY: BELONGS TO THE DETHIOBIOTIN SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE007647; AAK79329.1; -
 DR PIR; F97067; F97067.
 DR HAMAP; MF_00336; -; 1.
 DR InterPro; IPR004472; BiOD.
 DR InterPro; IPR002586; CblA_P.
 DR Pfam; PF01656; CblA_1.
 DR TIGRFAMs; TIGR00347; BiOD_1.
 KW Biotin biosynthesis; Ligase; Magnesium; ATP-binding;
 KW Complete proteome.
 KW NP_BIND 8
 KW SEQUENCE 240 AA; 26677 MW; 2F9303A8C3EAD1 CRC64;
 SQ
 Query Match 42.4%; Score 47.5; DB 1; Length 240;
 Best Local Similarity 45.8%; Pred. No. 2.2;
 Matches 11; Conservative 4; Mismatches 6; Indels 3; Gaps 1;
 QY 2 GARTINGOLL---YFRANGVOYKG 22
 DB 152 GIGTINHTVLTYSYKNGNIEYKG 175
 RESULT 10
 DOC_BPPI STANDARD; PRT; 126 AA.
 AC Q06259;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Death ON curing protein.
 GN DOC.
 OC Bacteriophage P1.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC P1-like viruses.
 OC NCBI_TaxID=10678;
 RX SEQUENCE FROM N.A.
 RX MEDLINE=94016561; PubMed=8411157;
 RA Lehner H., Maguin E., Jafri S., Yarmolinsky M.B.;

RT "Plasmid addiction genes of bacteriophage P1: doc, which causes cell
 RT death on curing of prophage, and phd, which prevents host death when
 RT prophage is retained."
 RL J. Mol. Biol. 233:414-428(1993).
 CC -1- FUNCTION: KILLS CELLS. DOC AND PHD PROTEIN FUNCTION IN UNISSON TO
 CC STABILIZE PLASMID NUMBER BY INDUCING A LETHAL RESPONSE TO PLASMID
 CC LOSS.
 CC -1- MISCELLANEOUS: THE CONCENTRATION OF PHD IN P1 LYSOGENS IS FAR
 CC GREATER THAN THAT OF THE POISON IT ANTAGONIZES. SUCH AN EXCESS MAY
 CC ASSUME THE WELL-BEING OF CARRIERS OF THE ADDICTING PLASMID.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M95666; AAA16931.1; -
 DR PIR; S40016; S40016.
 DR InterPro; IPR006440; DOC_P1.
 DR Pfam; PF05012; DOC; 1.
 DR TIGRFAMs; TIGR01550; DOC_P1_1.
 SQ SEQUENCE 126 AA; 13588 MW; E04A397538A595CD CRC64;
 Query Match 42.0%; Score 47; DB 1; Length 126;
 Best Local Similarity 60.0%; Pred. No. 1.4; Indels 0; Gaps 0;
 Matches 9; Conservative 2; Mismatches 4;
 QY 6 INGOLLYFRANGVOY 20
 DB 77 LNSHLLFLRNGVOY 91
 RESULT 11
 SPL_MOUSE STANDARD; PRT; 781 AA.
 ID SPL_MOUSE
 AC O89090; Q62251; Q64167;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transcription factor Spl.
 GN Spl.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RX SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Neuroblastoma;
 RX MEDLINE=98290554; PubMed=9628590;
 RA Yajima S., Lee S.H., Minowa T., Mouradian M.M.;
 RT "Sp family transcription factors regulate expression of rat D2
 RT dopamine receptor gene."
 RT DNA Cell Biol. 17:471-479(1998).
 RL (12)
 RN RN
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RP MEDLINE=96016118; PubMed=7568082;
 RA Peregrine S.P., Saffer J.D., Kilpatrick D.L.;
 RT "An alternatively spliced form of the transcription factor Spl
 RT containing only a single glutamine-rich transactivation domain."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9107-9111(1995).
 RN (13)
 RP SEQUENCE OF 681-781 FROM N.A.
 RP MEDLINE=92338398; PubMed=1633330;
 RA Chestier A., Charney P.;
 RT "Difference in the genomic organizations of the related transcription
 RT factors Spl and Krox-20: possible evolutionary significance."
 RL DNA Seq. 2:325-327(1992).
 CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
 CC RECOGNITION SITES.

```

CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=089090-1; Sequence=Displayed;
CC Name=2; Synonyms=Spl-5;
CC IsoId=089090-2; Sequence=VSP 007376;
CC -1- PTM: O-GLYCOSYLATED; CONTRAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF062566; AAC16484.1; -
CC EMBL; S79832; AAB35321.1; -
CC EMBL; X60136; CAA42721.1; -
CC HSPB; P08047; 1SP1.
CC MGD; MGI:98372; Spl.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0003700; F:transcription factor; IMP.
CC GO; GO:000355; P:regulation of transcription, DNA-dependent; IMP.
CC InterPro: IPR007087; Znf.C2H2.
CC Pfam: PF00096; zfc-C2H2; 3.
CC ProDom: PD000003; Znf.C2H2; 2.
CC SMART; SM00355; Znf.C2H2; 3.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
CC Transcription regulation; Activator; Zinc-finger; Metal-binding;
CC DNA-binding; Nuclear protein; Repeat; Glycoprotein;
CC KW Alternative splicing
CC FT ZN_FING 624 648 C2H2-TYPE 1.
CC FT ZN_FING 654 678 C2H2-TYPE 2.
CC FT ZN_FING 684 706 C2H2-TYPE 3.
CC FT VARSPLIC 57 370 Missing (in isoform 2).
CC FT VARSPLIC /FridaVSP 007376.
CC FT VARSPLIC V -> G (in REF. 2).
CC FT CONFLICT 459 459
CC SEQENCE 781 AA; 80486 MW; 14CD12B8C58CF21 CRC64;
SQ

```

```

Query Match 42.0%; Score 47; DB 1; Length 781;
Best Local Similarity 50.0%; Pred. No. 8.9;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 RTINGQLLFRANGVQYK 21
Db 179 QTVDGQQLQPAATGAQVQ 196

```

```

RESULT 12
SPL_HUMAN STANDARD; PRT; 785 AA.
AC P08047; Q9H305; Q9NR51; Q9NY21; Q9NYE7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor Spl.
DE SPL OR TSEF1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE OF 4-785 FROM N.A.
RC TISSUE=Cervical carcinoma;
RA Haggart M.H., Ladurner A.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE OF 1-558 FROM N.A.
RX MEDLINE=20545561; PubMed=10973950;
RA Takahara T., Kanazu S., Yanagisawa S., Akenawa H.,
RT "Heterogeneous Spl mRNAs in human HepG2 cells include a product of
RL homotypic trans-splicing."
RL J. Biol. Chem. 275:38067-38072(2000).
RN [3]
RP SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675.
RX MEDLINE=8080466; PubMed=3319186;
RA Kadonaga J.T., Garner K.R., Masiarz F.R., Tjian R.,
RT "Isolation of cDNA encoding transcription factor Spl and functional
RL analysis of the DNA binding domain."
RL Cell 51:1079-1090(1987).
RN [4]
RP SEQUENCE OF 1-109 FROM N.A.
RA Nicolas M., Noe V., Ciudad C.J.,
RT "Expression of transcription factor Spl mRNA in mammalian cells."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-98 FROM N.A.
RA Handebug K., Huebner A.,
RT "Sequencing of the 5' end of human transcription factor Spl mRNA."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP O-GLYCOSYLATION.
RX MEDLINE=8903041; PubMed=3139301;
RA Jackson S.P., Tjian R.,
RT "O-glycosylation of eukaryotic transcription factors: implications
RL for mechanisms of transcriptional regulation."
RL Cell 55:125-133(1988).
RN [7]
RP STRUCTURE BY NMR OF 654-684 AND 684-712.
RX MEDLINE=97218212; PubMed=9065444;
RA Narayan V.A., Kiriacki R.W., Caradonna J.P.,
RT "Structures of zinc finger domains from transcription factor Spl.
RL Insights into sequence-specific protein-DNA recognition."
RL J. Biol. Chem. 272:7801-7809(1997).
RN [8]
RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
RX MEDLINE=96224025; PubMed=8626793;
RA Parks C.L., Shenk T.,
RT "The serotonin 1a receptor gene contains a TATA-less promoter that
RL responds to MAZ and Spl."
RL J. Biol. Chem. 271:4417-4430(1996).
RN [9]
RP FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
CC SEROTONIN RECEPTOR PROMOTER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: O-GLYCOSYLATED; CONTRAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
CC -1- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF252284; AAF67726.1; -
CC EMBL; AB039286; BAB13476.1; -
CC EMBL; J03133; AAA61154.1; -
CC EMBL; AF255682; AAF78781.1; -
CC EMBL; AJ272134; CAB75345.1; -
CC PIR; A29635; A29635.
CC PDB; 1SP1; 21-APR-97.
CC PDB; 1SP2; 21-APR-97.
CC TRANSPAC; T00759; -
CC GlycoSiteDB; P08047; -

```

```

DR Genew; HGNC:11205; SPL.
DR MTM; 189905; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; NAS.
DR GO; GO:0016563; F:transcriptional activator activity; NAS.
DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; NAS.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
DR Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
FT ZN_FING 626 650
FT ZN_FING 656 680 C2H2-TYPE 2.
FT ZN_FING 686 708 C2H2-TYPE 3.
FT CONFLICT 366 366 D->G (IN REF. 3; AA SEQUENCE).
FT CONFLICT 670 670 S->F (IN REF. 3; AA SEQUENCE).
FT STRAND 657 657
FT STRAND 661 662
FT STRAND 666 666
FT HELIX 670 677
FT TURN 678 680
SQ SEQUENCE 785 AA; 80693 MW; 43893DBF6518B9EA CRC64;

Query Match 42.0%; Score 47; DB 1; Length 785;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 RTINGQLLYFRANGVQYK 21
Db 180 QTVDGQQLQFPATGAQVQ 197

RESULT 13
SPL_RAT
ID SPL_RAT STANDARD; PRT; 788 AA.
AC Q01714;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor Spl.
GN Spl.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93010958; PubMed=1356762; Kikuchi Y., Sasano K.,
RA Imataka H., Sogawa K., Yasumoto K., Fujii-Kuriyama Y.;
RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
RT "Two regulatory proteins that bind to the basic transcription element
(BME), a GC box sequence in the promoter region of the rat P-4501A1
gene."
RL EMBL J. 11:3663-3671 (1992).
CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; D12768; BAA02235.1; -.
DR PIR; J05747; J05747.
DR HSSP; P08047; ISPL.
DR TRANSFAC; T00754; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
DR Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein.
FT ZN_FING 629 653
FT ZN_FING 659 683 C2H2-TYPE 1.
FT ZN_FING 689 711 C2H2-TYPE 3.
SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;

Query Match 42.0%; Score 47; DB 1; Length 788;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 RTINGQLLYFRANGVQYK 21
Db 183 QTVDGQQLQFPATGAQVQ 200

RESULT 14
ID DLH AQUAE STANDARD; PRT; 231 AA.
AC 067802;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative carboxymethyl-enebutenolidease (EC 3.1.1.45) (Dienelactone
DE hydrolase) (DLH).
GN AQ 1997.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxId=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPS;
RX MEDLINE=9819666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Ausley M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."
RL Nature 392:353-358 (1998).
CC -1- CATALYTIC ACTIVITY: 4-carboxymethyl-enebut-2-en-4-olide + H(2)O = 4
CC oxohe-2-enedioate.
CC -1- SIMILARITY: BELONGS TO THE DIENELACTONE HYDROLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000767; AAC07773.1; -.
DR PIR; F70471; F70471.
DR InterPro; IPR002925; DLH.
DR InterPro; IPR00379; Ser_estr_site.
DR Pfam; PF01738; DLH; 1.
KW Hypothetical protein; Hydrolase; Complete proteome.
FT ACT_SITE 118 118 BY SIMILARITY.
FT ACT_SITE 167 167 BY SIMILARITY.
FT ACT_SITE 199 199 BY SIMILARITY.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 33.1564 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-3
Perfect score: 112
Sequence: 1 TGARTINGQILYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.4	591	2	Q8VUH3	Q8VUH3 streptococc
2	87	1590	2	Q59983	Q59983 streptococc
3	82	1599	2	Q00599	Q00599 streptococc
4	81	72.3	2	Q8G9Q2	Q8G9Q2 leucococc
5	80	71.4	2	Q00600	Q00600 streptococc
6	76	1590	2	Q55263	Q55263 streptococc
7	75	67.0	2	Q8DUM9	Q8DUM9 streptococc
8	75	67.0	2	Q8756	Q8756 leucococc
9	75	1290	2	Q8WU4	Q8WU4 streptococc
10	73	65.2	2	Q5264	Q5264 streptococc
11	72	64.3	2	Q91466	Q91466 leucococc
12	72	64.3	2	Q9EZH5	Q9EZH5 leucococc
13	72	64.3	2	Q52224	Q52224 leucococc
14	72	64.3	2	Q54178	Q54178 streptococc
15	68	60.7	2	Q54447	Q54447 streptococc
16	68	60.7	2	Q8DRV2	Q8DRV2 streptococc

17	68	60.7	1449	2	Q68542	Q68542 streptococc
18	66	58.9	1577	2	Q55265	Q55265 streptococc
19	65.5	58.5	330	2	Q55228	Q55228 streptococc
20	65	58.0	2057	2	Q9RE05	Q9RE05 leucococc
21	64	57.1	1512	2	Q9WUJ5	Q9WUJ5 streptococc
22	60.5	54.0	1554	2	Q8KZL5	Q8KZL5 streptococc
23	57	50.9	1527	2	Q9ZAR4	Q9ZAR4 leucococc
24	57	50.9	1527	2	Q8KRE1	Q8KRE1 leucococc
25	57	50.9	1575	2	Q9LCH3	Q9LCH3 streptococc
26	55	49.1	2817	16	Q97K42	Q97K42 clostridium
27	54	48.2	221	5	Q44353	Q44353 caenorhabdi
28	53	47.3	1002	10	Q8GZY5	Q8GZY5 oryza sativ
29	50	44.6	831	2	Q50076	Q50076 plasmodium
30	49	43.8	372	5	Q8TW25	Q8TW25 plasmodium
31	49	43.8	491	16	Q97JB2	Q97JB2 clostridium
32	47	42.0	126	2	Q47171	Q47171 escherichia
33	47	42.0	126	2	Q47173	Q47173 escherichia
34	47	42.0	126	2	Q47172	Q47172 escherichia
35	47	42.0	255	2	Q66378	Q66378 clostridium
36	47	42.0	396	2	Q9P932	Q9P932 clostridium
37	47	42.0	559	11	Q8K4R0	Q8K4R0 ratius norv
38	47	42.0	781	11	Q89090	Q89090 mus musculu
39	47	42.0	784	11	Q89087	Q89087 mus musculu
40	46	41.1	608	2	Q8VQ55	Q8VQ55 streptococc
41	46	41.1	619	2	Q54972	Q54972 streptococc
42	46	41.1	619	16	Q8DR10	Q8DR10 streptococc
43	46	41.1	744	16	Q97T39	Q97T39 streptococc
44	46	41.1	1301	2	Q06307	Q06307 alicyclobac
45	46	41.1	2364	2	Q46342	Q46342 clostridium

ALIGNMENTS

RESULT 1

ID	Q8VUH3	PRELIMINARY;	PRT;	591 AA.
AC	Q8VUH3	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	Glucosyltransferase SI (Fragment).			
GN	GTFC.			
OS	Streptococcus mutans.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OX	NCBI_TaxID=1309;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Xc; Yamashita Y., Shibata Y., Nakano Y., Koga T., Tada H.,			
RA	"Genes Involved in Bacitracin Resistance in Streptococcus mutans.",			
RT	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
RL	EMBL: AB078507; BAB83942.1; -			
DR	InterPro: IPR002479; CM_binding.			
DR	InterPro: IPR003318; Glyco_hydro_70.			
DR	Pfam: PF01473; CM_binding_1; 9.			
DR	Pfam: PF02324; Glyco_hydro_70; 1.			
KM	Transferase.			
FT	NON TER			
SO	SEQUENCE	1	1	
	591 AA;	67094 MW;	0933DCE4421DAF30 CRC64;	

Query Match 80.4%; Score 90; DB 2; Length 591;
Best local Similarity 72.7%; Pred. No. 4.6e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQILYFRANGVQVKG 22
DB 465 TGARTINGQILYFRANGVQVKG 486

RESULT 2
Q59983

ID 059983 PRELIMINARY; PRT; 1590 AA.
AC 059983;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5).
GN GTFI.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OM2176;
RA MEDLINE=94146405; PubMed=8112602;
SA Sato S., Inoue M., Handa N., Aizawa Y., Isohe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
ST Streptococcus sobrinus.";
RL DNA Seq. 4:19-27(1993).
DR EMBL; D13858; BAA02976.1; -;
DR InterPro; IPR002479; CM binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Glycosyltransferase; Signal; Transferase.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 1 1590 GLUCOSYLTRANSFERASE-I.
SQ SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 77.7%; Score 87; DB 2; Length 1590;
Best Local Similarity 77.3%; Pred. No. 4.4e-05;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGOVKG 22
DB 1296 TGAQTIRGQKLYFKANGQOVKG 1317

RESULT 3
Q00599 PRELIMINARY; PRT; 1599 AA.
ID 000599;
AC 000599;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Glucosyltransferase S precursor (EC 2.4.1.5) (GTF) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFK.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25975;
RA MEDLINE=93381463; PubMed=8371114;
RX Giffard P.M., Allen D.M., Milward C.P., Simpson C.L., Jacques N.A.;
RT "Sequence of the gtfK gene of Streptococcus salivarius ATCC 25975 and
evolution of the gtf genes of oral streptococci.";
RT U. Gen. Microbiol. 139:1511-1522(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25975;
RA MEDLINE=92148377; PubMed=1838391;
RX Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
RT "Molecular characterization of a cluster of at least two
glucosyltransferase genes in Streptococcus salivarius ATCC 25975.";
RT J. Gen. Microbiol. 137:2577-2593(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-

CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- DISEASE: DENTAL CARIES.
CC -1- SIMILARITY: TO REGIONS OF BARLEY AND BACILLUS AMYLOLIQUEPACIENS
CC ALPHA AMYLASES AND RABBIT GLYCOGEN PHOSPHORYLASE.
DR EMBL; Z11873; CAA77898.1; -;
DR EMBL; Z11873; CAA77901.1; -;
DR EMBL; M6111; AAA26887.1; -;
DR InterPro; IPR002479; CM_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 42 POTENTIAL.
FT CHAIN 1 1599 GLUCOSYLTRANSFERASE S.
SQ SEQUENCE 1599 AA; 176480 MW; 24B77869P152B707 CRC64;

QY 1 TGARTINGOLLYFRANGOVKG 22
DB 1402 TGAQVINGOHLYPDANGOVKG 1423

RESULT 4
Q08902 PRELIMINARY; PRT; 2835 AA.
ID 08902;
AC 08902;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Dextranucrase (EC 2.4.1.5) (Fragment).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22231661; PubMed=12270834;
RX Bozonnet S., Dols-Lafargue M., Fabre E., Fizzut S., Renaud-Simeon M.,
Mondan P., Willemot R.M.;
RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
synthesising dextranucrase with two catalytic domains.";
RT J. Bacteriol. 184:5753-5761(2002).
RL EMBL; AJ430204; CAD22883.1; -;
KM Transferase; Glycosyltransferase.
FT NON TER 1 1
SQ SEQUENCE 2835 AA; 313264 MW; D03262CDJ735399D CRC64;

Query Match 72.3%; Score 81; DB 2; Length 2835;
Best Local Similarity 72.7%; Pred. No. 0.00082;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGOVKG 22
DB 1644 TGAQTIRGQKLYFKANGQOVKG 1665

RESULT 5
Q00600 PRELIMINARY; PRT; 1518 AA.
ID 000600;
AC 000600;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Glucosyltransferase I (EC 2.4.1.5) (GTF) (Dextranucrase) (Sucrose 6-
glucosyltransferase).
GN GTFU.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.

```

OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25975;
RX MEDLINE=92148377; PubMed=1838391;
RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
RT "Molecular characterization of a cluster of at least two
RT glucosyltransferase genes in Streptococcus salivarius ATCC 25975."
RT J. Gen. Microbiol. 137:2577-2593(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- DISEASE: DENTAL CARIES.
DR EMBL: Z11873; CAA77900.1; -.
DR EMBL: M64111; AAA26896.1; -.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW binding 1; 13.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KM Transferrase: Glycosyltransferase; Repeat; Dental caries.
FT DOMAIN 1307 1482 6 DIRECT REPEATS.
FT REPEAT 1307 1338 PEPEAT 1.
FT REPEAT 1339 1352 PEPEAT 2.
FT REPEAT 1372 1403 PEPEAT 3.
FT REPEAT 1404 1417 PEPEAT 4.
FT REPEAT 1437 1468 PEPEAT 5.
FT REPEAT 1469 1482 PEPEAT 6.
SQ SEQUENCE 1518 AA; 167730 MW; DAA41F717098B59A CRC64;

Query March 71.4%; Score 80; DB 2; Length 1518;
Best Local Similarity 72.7%; Pred. No. 0.0058;
Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVK 22
DB 1383 TGAQTVGKQLLYFRANGVQVK 1404

RESULT 6
Q55263 PRELIMINARY; PRT; 1590 AA.
ID Q55263
AC Q55263;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 05, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE GTF-1.
OS GLUCOSYLTRANSFERASE.
OC Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glycan synthetase
RT produced from Streptococcus sobrinus ATCC 33478."
RT Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL: D63570; BAA09792.1; -.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW binding 1; 15.
DR Pfam: PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3FC6BFAFD43 CRC64;

Query March 67.9%; Score 76; DB 2; Length 1590;
Best Local Similarity 68.2%; Pred. No. 0.0028;
Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 TGARTINGQLLYFRANGVQVK 22
DB 1231 TGAQTVGKQLLYFRANGVQVK 1252

RESULT 7
Q8DUW9 PRELIMINARY; PRT; 726 AA.
ID Q8DUW9
AC Q8DUW9;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Putative glycan-binding protein D, BglB-like protein.
GN GBP OR SMU.772.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22285063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson W.B., Primeaux C., Tian R., Xenon S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen."
RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RL EMBL: AB014918; BAA58492.1; -.
KM Complete proteome.
SQ SEQUENCE 726 AA; 79787 MW; B2AB7C06F6E7CBAE CRC64;

Query March 67.0%; Score 75; DB 16; Length 726;
Best Local Similarity 63.6%; Pred. No. 0.0016;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVK 22
DB 235 TGSITVNGSLYFRANGVQVK 256

RESULT 8
Q48756 PRELIMINARY; PRT; 1290 AA.
ID Q48756
AC Q48756;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Dextranucrase.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B1299;
RX MEDLINE=97136686; PubMed=8982063;
RA Monchois V., Willemot R.M., Renaud-Simeon M., Croux C., Monsen P.;
RT "Cloning and sequencing of a gene coding for a novel dextranucrase
RT from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-
RT 6) and alpha (1-3) linkages."
RT Gene 182:23-32(1996).
RL EMBL: U38181; AAA40875.1; -.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW binding 1; 11.
DR Pfam: PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1290 AA; 145590 MW; 3555C2B96B749F9AA CRC64;

Query March 67.0%; Score 75; DB 2; Length 1290;
Best Local Similarity 68.2%; Pred. No. 0.0032;
Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```


Db 1245 TGVVINGQTLTFPDAGRVKG 1266

RESULT 9

Q9WKJ4 PRELIMINARY; PRT; 1338 AA.
AC Q9WKJ4
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE GTF-S.
GN GTF-S.
OS Streptococcus criceti.
OC Plasmid PAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S. cricetus glucosyltransferase (gifs and gift) genes."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026123; BAA77236.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90CBE10E15D99B CRC64;

Query Match 67.0%; Score 75; DB 2; Length 1338;
Best Local Similarity 63.6%; Pred. No. 0.0033;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVG 22
Db 1207 TGSQTAGQXVFQPNQVQVG 1228

RESULT 10

Q55264 PRELIMINARY; PRT; 1449 AA.
AC Q55264
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTF.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122197; Pubmed=7822030;
RA Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL; L35495; AAC41412.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Signal; Transferase.
FT SIGNAL 1
FT CHAIN 35
SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 65.2%; Score 73; DB 2; Length 1449;
Best Local Similarity 68.2%; Pred. No. 0.0077;

Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVG 22
Db 1258 TGHQNINGQELTFDNNQVQVG 1279

RESULT 11

Q9L466 PRELIMINARY; PRT; 1477 AA.
AC Q9L466
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRG.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Renaud-Simeon K., Pizzut S., Sarcabal P.,
RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250172; CAB76565.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
KW Dextranucrase.
SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFC8B31 CRC64;

Query Match 64.3%; Score 72; DB 2; Length 1477;
Best Local Similarity 68.2%; Pred. No. 0.011;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVG 22
Db 1368 TGLVNINGMLKXFQANGRVQVG 1389

RESULT 12

Q9EZHS PRELIMINARY; PRT; 1508 AA.
AC Q9EZHS
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP294469; AAG38021.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Signal; Transferase.
FT SIGNAL 1
FT CHAIN 1508 AA; 168542 MW; E2FCEA0F97AE4F3A CRC64;
SQ SEQUENCE 1508 AA; 168542 MW; E2FCEA0F97AE4F3A CRC64;

Query Match 64.3%; Score 72; DB 2; Length 1508;
Best Local Similarity 68.2%; Pred. No. 0.012;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TGARTINGQLLYFRANGVQYKG 22
 Db 1399 TGLVNINGNLTQFQANGROVKG 1420

RESULT 13
 052224 PRELIMINARY; PRT; 1508 AA.

ID 052224
 AC 052224
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Renaud-Simeon M., Monsan P., Millemot R.M.;
 RT "Cloning and sequencing of a gene coding for an extracellular
 RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
 RT synthesizing only a a(1-6) glucan."
 RL FEWS Microbiol. Lett. 0:0-0(1998).
 DR EMBL; AF030129; AAB95453.1;
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003118; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 16851 MW; E70CECB57A70D1F0 CRC64;

Query Match 64.3%; Score 72; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.012;
 Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TGARTINGQLLYFRANGVQYKG 22
 Db 1399 TGLVNINGNLTQFQANGROVKG 1420

RESULT 14
 054178 PRELIMINARY; PRT; 1577 AA.

ID 054178
 AC 054178; Q54247; 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase.
 GN GTFG.
 OS Streptococcus gordonii Challis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=293390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHALLIS;
 RX MEDLINE=96157084; PubMed=8586195;
 RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
 RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
 RT phase variants."
 RL Dev. Biol. Stand. 85:309-314(1995).
 RN [2]
 RP SEQUENCE OF 1-96 FROM N.A.
 RC STRAIN=CHALLIS;
 RX MEDLINE=92276337; PubMed=1534336;
 RA Sulavik M.C., Tardif G., Clewell D.B.;
 RT "Identification of a gene, rgg, which regulates expression of
 RT glucosyltransferase and influences the Spp phenotype of Streptococcus
 RT gordonii Challis."
 RL J. Bacteriol. 174:3577-3586(1992).
 DR EMBL; U12643; AAC3483.1; -.

DR EMBL; M89776; AAA26969.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003118; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 18.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase.
 SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match 64.3%; Score 72; DB 2; Length 1577;
 Best Local Similarity 63.6%; Pred. No. 0.012;
 Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TGARTINGQLLYFRANGVQYKG 22
 Db 1379 TGLKTIINVLYPFDQKQVKG 1400

RESULT 15
 054447 PRELIMINARY; PRT; 563 AA.

ID 054447
 AC 054447
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucan-binding protein (gbp).
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Banas J.A., Russell R.R.B., Ferretti J.J.;
 RT "Sequence analysis of the gene for the glucan-binding protein of
 RT Streptococcus mutans INGBRITT."
 RL Submitted (MAY-1990) to the EMBL/genbank/DBJ databases.
 DR EMBL; M30945; AAA26894.1; -
 DR InterPro; IPR002479; CW binding.
 DR Pfam; PF01473; CW_binding_1; 9.
 KW Streptococcus.
 SQ SEQUENCE 563 AA; 62932 MW; A0BB6DAF42D1B5F CRC64;

Query Match 60.7%; Score 68; DB 2; Length 563;
 Best Local Similarity 66.7%; Pred. No. 0.017;
 Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GARTINGQLLYFRANGVQYKG 22
 Db 276 GMRITGGKXYFDITNGVQYKG 296

Search completed: November 13, 2003, 09:44:01
 Job time : 35.1564 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 14.2844 Seconds
(Without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-3
Perfect score: 112
Sequence: 1 TGARTINGQLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PTCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	3	US-09-007-999-2
2	112	100.0	1475	3	US-09-210-361-2
3	112	100.0	1475	4	US-09-740-274-2
4	87	77.7	22	1	US-08-057-162B-3
5	78	63.6	1375	3	US-09-210-361-4
6	78	63.6	1375	4	US-09-740-274-4
7	67	59.8	1430	3	US-09-008-172-2
8	67	59.8	1430	3	US-09-210-361-6
9	67	59.8	1430	4	US-09-740-274-6
10	66	58.9	1577	2	US-08-793-824-2
11	65	58.0	2057	4	US-09-499-203-2
12	49.5	44.2	811	1	US-08-480-604A-7
13	49.5	44.2	811	2	US-08-405-496A-7
14	49.5	44.2	811	3	US-08-915-136-7
15	49.5	44.2	811	4	US-08-957-310-7
16	49.5	44.2	811	4	US-10-011-366-7
17	49.5	44.2	812	1	US-08-480-604A-29
18	49.5	44.2	812	3	US-08-915-136-23
19	49.5	44.2	2710	1	US-08-480-604A-6
20	49.5	44.2	2710	2	US-08-405-496A-6
21	49.5	44.2	2710	3	US-08-915-136-6
22	49.5	44.2	2710	4	US-08-957-310-6
23	49.5	44.2	2710	4	US-10-011-366-6
24	46	41.1	619	1	US-08-465-746-2
25	46	41.1	619	1	US-08-214-164-2
26	46	41.1	619	2	US-08-467-852A-3
27	46	41.1	619	2	US-08-246-636-2

28	46	41.1	619	2	US-08-247-491A-3	Sequence 3, Appl
29	46	41.1	619	2	US-08-319-795-2	Sequence 2, Appl
30	46	41.1	619	2	US-08-468-985-2	Sequence 2, Appl
31	46	41.1	619	3	US-08-312-949-2	Sequence 2, Appl
32	46	41.1	641	3	US-08-961-083-160	Sequence 160, App
33	46	41.1	641	4	US-09-536-784-160	Sequence 2, Appl
34	46	41.1	648	1	US-08-072-070-2	Sequence 2, Appl
35	46	41.1	648	1	US-08-469-434-2	Sequence 2, Appl
36	46	41.1	648	1	US-08-214-222-2	Sequence 2, Appl
37	46	41.1	648	2	US-08-467-852A-2	Sequence 2, Appl
38	46	41.1	648	2	US-08-468-718-2	Sequence 2, Appl
39	46	41.1	648	3	US-08-247-491A-2	Sequence 3, Appl
40	46	41.1	648	3	US-08-446-201-3	Sequence 23, Appl
41	46	41.1	695	1	US-08-127-499A-23	Sequence 23, Appl
42	46	41.1	695	1	US-08-482-847-23	Sequence 41, Appl
43	46	41.1	1231	4	US-08-714-741-41	Sequence 2, Appl
44	43	38.4	153	4	US-09-287-070-2	Sequence 73, Appl
45	43	38.4	1036	4	US-09-206-942-73	

ALIGNMENTS

```

RESULT 1
US-09-007-999-2
Sequence 2, Application US/09007999
Patent No. 6087559
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: 0356D
CURRENT APPLICATION NUMBER: US/09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match      100.0%; Score 112; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 TGARTINGQLYFRANGVQVKG 22
DB      1300 TGARTINGQLYFRANGVQVKG 1321

RESULT 2
US-09-210-361-2
Sequence 2, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
EARLIER FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-210-361-2

Query Match          100.0%; Score 112; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGARTINGQLLYFRANGVQYKG 22
Db      1300 TGARTINGQLLYFRANGVQYKG 1321

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-2

Query Match          100.0%; Score 112; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGARTINGQLLYFRANGVQYKG 22
Db      1300 TGARTINGQLLYFRANGVQYKG 1321

RESULT 4
US-08-057-162B-3
; Sequence 3, Application US/08057162B
; Patent No. 5686075
; GENERAL INFORMATION:
; APPLICANT: Taubman, Martin A.
; APPLICANT: Smith, Daniel J.
; TITLE OF INVENTION: SYNTHETIC PEPTIDE VACCINE FOR DENTAL CARIES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
```

```

; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/057,162B
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/877,295
; FILING DATE: 01-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wagner, Richard W.
; REGISTRATION NUMBER: 34,480
; REFERENCE/DOCKET NUMBER: FDC92-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-057-162B-3

Query Match          77.7%; Score 87; DB 1; Length 22;
Best Local Similarity 77.3%; Pred. No. 1.9e-08;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 TGARTINGQLLYFRANGVQYKG 22
Db      1 TGAQTINGQLLYFRANGVQYKG 22

RESULT 5
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT FILING DATE: 1998-12-11
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-210-361-4

Query Match          69.6%; Score 78; DB 3; Length 1375;
```

Best Local Similarity 68.2%; Pred. No. 9.9e-05;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22

Db 1264 TGTVTNGQRLYFKPNGVQAKG 1285

RESULT 6
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 69.6%; Score 78; DB 4; Length 1375;
Best Local Similarity 68.2%; Pred. No. 9.9e-05;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22

Db 1264 TGTVTNGQRLYFKPNGVQAKG 1285

RESULT 7
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 59.8%; Score 67; DB 3; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.0081;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22

Db 1332 TGSQTLAGKLYFASDGQVKG 1353

RESULT 8
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 59.8%; Score 67; DB 3; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.0081;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22

Db 1332 TGSQTLAGKLYFASDGQVKG 1353

RESULT 9
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6

LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 59.8%; Score 67; DB 4; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.0081;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1332 TGSQTIAKKTLYFASDQVQVKG 1353

RESULT 10
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of plants to
TITLE OF INVENTION: Increase Stored Carbohydrates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Giffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 58.9%; Score 66; DB 2; Length 1577;
Best Local Similarity 54.5%; Pred. No. 0.014;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1469 TGLQNNKYYFSGNQAQVQVKG 1490

RESULT 11
US-09-499-203-2
Sequence 2, Application US/09499203
Patent No. 6570065
GENERAL INFORMATION:

APPLICANT: KOSSMANN, Jens
APPLICANT: WELSH, Thomas
APPLICANT: QUANZ, Martin
APPLICANT: KNUTH, Karola
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
FILE REFERENCE: 147-196P
CURRENT APPLICATION NUMBER: US/09/499,203
CURRENT FILING DATE: 2000-02-08
NUMBER OF SEQ. ID NOS: 54
SOFTWARE: Patentin Ver. 2.1
SEQ. ID NO 2
LENGTH: 2057
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 58.0%; Score 65; DB 4; Length 2057;
Best Local Similarity 57.1%; Pred. No. 0.028;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GARTINGQLLYFRANGVQVKG 22
DB 261 GLQTTIDNLTQYFNGQVQVQVKG 281

RESULT 12
US-08-480-604A-7
Sequence 7, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCH, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-604A-7

Query Match 44.2%; Score 49.5; DB 1; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGQLLYFRAN-GVQVKG 22
Db 587 TGLRTIDGKKYENTNTAVAVTG 609

RESULT 13
US-08-405-496A-7
Sequence 7, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-405-496A-7

Query Match 44.2%; Score 49.5; DB 2; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGQLLYFRAN-GVQVKG 22
Db 587 TGLRTIDGKKYENTNTAVAVTG 609

RESULT 14
US-08-915-136-7
Sequence 7, Application US/08915136
Patent No. 6290960
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: protein
US-08-915-136-7

Query March 44.2%; Score 49.5; DB 3; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGOLLYFRAN-GVQVKG 22
Db 587 TGLRTIDGKKYFNTNTAVAVTG 609

RESULT 15
US-08-957-310-7

Sequence 7, Application US/08957310
Patent No. 6365158

GENERAL INFORMATION:

APPLICANT: Williams, James A.

TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES

TITLE OF INVENTION: OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE

TITLE OF INVENTION: DISEASE

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSES: Medlen & Carroll

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,310

FILING DATE: 23-OCT-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 24-OCT-1994

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPHD-01121

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 811 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-957-310-7

Search completed: November 13, 2003, 09:47:58
Job time: 15.2844 secs

Query Match 44.2%; Score 49.5; DB 4; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGOLLYFRAN-GVQVKG 22
Db 587 TGLRTIDGKKYFNTNTAVAVTG 609

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using bw model

Run on: November 13, 2003, 09:45:40 / Search time 28.7773 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-3
Perfect score: 112
Sequence: 1 TGARTINGQLTYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications AA:
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	9 US-09-740-274-2	Sequence 2, Appli
2	78	69.6	1375	9 US-09-740-274-4	Sequence 4, Appli
3	67	59.8	1430	9 US-09-740-274-6	Sequence 6, Appli
4	49.5	44.2	811	15 US-10-011-366-7	Sequence 7, Appli
5	49.5	44.2	866	16 US-10-222-038-2	Sequence 2, Appli
6	49.5	44.2	2710	15 US-10-011-366-6	Sequence 6, Appli
7	49	43.8	343	15 US-10-275-360-13	Sequence 13, Appli
8	48	42.9	120	15 US-10-156-761-13796	Sequence 13796, A
9	46	41.1	619	11 US-09-882-774-1	Sequence 1, Appli
10	46	41.1	641	9 US-09-765-272-160	Sequence 160, App
11	45	41.1	744	11 US-09-769-787-184	Sequence 184, App
12	44	39.3	115	9 US-09-764-869-970	Sequence 970, App
13	44	39.3	115	15 US-10-091-504-970	Sequence 970, App
14	43.5	38.8	196	10 US-09-738-626-4344	Sequence 4344, Ap
15	43	38.4	153	9 US-09-287-070-2	Sequence 2, Appli

16	43	38.4	153	15 US-10-243-977-2	Sequence 2, Appli
17	43	38.4	349	12 US-10-274-694-13	Sequence 13, Appli
18	43	38.4	1036	12 US-10-193-764-69	Sequence 69, Appli
19	43	38.4	1477	12 US-10-193-764-67	Sequence 67, Appli
20	43	38.4	1477	14 US-10-092-880-4	Sequence 4, Appli
21	42	37.5	165	15 US-10-156-761-14299	Sequence 14299, A
22	42	37.5	284	15 US-10-156-761-10979	Sequence 10979, A
23	42	37.5	929	11 US-09-298-523B-60	Sequence 60, Appli
24	41.5	37.1	472	12 US-10-029-386-33685	Sequence 33685, A
25	41.5	37.1	472	12 US-09-826-509-469	Sequence 469, App
26	41.5	37.1	472	15 US-10-225-567A-84	Sequence 49, Appli
27	41	36.6	412	12 US-10-301-997-49	Sequence 49, Appli
28	41	36.6	412	12 US-10-301-997-87	Sequence 87, Appli
29	41	36.6	678	10 US-09-738-626-4369	Sequence 4369, Ap
30	40	35.7	400	10 US-09-738-626-4358	Sequence 4358, Ap
31	40	35.7	499	15 US-10-156-761-11375	Sequence 11375, A
32	40	35.7	673	15 US-10-294-561-3	Sequence 3, Appli
33	39.5	35.3	251	12 US-10-366-686-6	Sequence 8, Appli
34	39	34.8	91	15 US-10-011-366-8	Sequence 1, Appli
35	39	34.8	179	9 US-09-287-070-1	Sequence 1, Appli
36	39	34.8	179	15 US-10-243-977-1	Sequence 1, Appli
37	39	34.8	232	15 US-10-198-070-34	Sequence 34, Appli
38	39	34.8	345	15 US-10-128-714-3396	Sequence 3396, Ap
39	39	34.8	356	15 US-10-128-714-8396	Sequence 8396, Ap
40	39	34.8	372	15 US-10-177-293-348	Sequence 348, App
41	39	34.8	453	9 US-09-765-272-38	Sequence 38, Appli
42	39	34.8	539	11 US-09-298-523B-54	Sequence 54, Appli
43	39	34.8	631	12 US-09-829-382-25	Sequence 25, Appli
44	39	34.8	655	11 US-09-056-019-2	Sequence 2, Appli
45	39	34.8	663	11 US-09-298-523B-58	Sequence 58, Appli

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 100.0%; Score 112; DB 9; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGARTINGQLTYFRANGVQVKG 22
|||||

Thu Nov 13 12:12:20 2003

us-09-290-049a-3.rapb

Page 2

Db 1300 TGAATNGQLLYFRANGVQYK 1321

RESULT 2

US-09-740-274-4
Sequence 4, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 69.6%; Score 79; DB 9; Length 1375;
Best Local Similarity 68.2%; Pred. No. 0.00022;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGAATNGQLLYFRANGVQYK 22

Db 1264 TGTVTNGQLLYFRANGVQYK 1285

RESULT 3

US-09-740-274-6
Sequence 6, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT

ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 59.8%; Score 67; DB 9; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.018;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGAATNGQLLYFRANGVQYK 22

Db 1332 TGSQTLAGKLYFRASDEQYK 1353

RESULT 4

US-10-011-366-7
Sequence 7, Application US/10011366
Publication No. US20030054493A1
GENERAL INFORMATION:
APPLICANT: Williams, James A.
Klink, John A.
TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
DISEASE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESS: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/011,366
FILING DATE: 16-NO. US20030054493A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/957,310
FILING DATE: 23-OCT-1997
APPLICATION NUMBER: US 08/329,154
FILING DATE: 24-OCT-1994
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-011-366-7

Query Match 44.2%; Score 49.5; DB 15; Length 811;
Best Local Similarity 52.2%; Pred. No. 9;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGAATNGQLLYFRANGVQYK 22

DB 587 TGLRTIDGKKYFNTNTAVATG 609

RESULT 5

US-10-222-038-2
Sequence 2, Application US/10222038
Publication No. US20030129198A1

GENERAL INFORMATION:

APPLICANT: Techlabs, Inc.

APPLICANT: Wilkins, Tracy

APPLICANT: Lyster, David

APPLICANT: Moncrief, Scott

APPLICANT: Zheng, Lamin

APPLICANT: Phelps, Carol

TITLE OF INVENTION: METHODS TO PRODUCE HIGH LEVELS OF C. DIFFICILE TOXINS

FILE REFERENCE: 42052-20002.10

CURRENT APPLICATION NUMBER: US/10/222,038

CURRENT FILING DATE: 2002-08-15

PRIOR APPLICATION NUMBER: US 09/545,773

PRIOR FILING DATE: 2000-04-10

PRIOR APPLICATION NUMBER: US 60/128,686

PRIOR FILING DATE: 1999-04-09

PRIOR APPLICATION NUMBER: US 60/186,201

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,111

PRIOR FILING DATE: 2000-03-20

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 866

TYPE: PRT

ORGANISM: Clostridium difficile

US-10-222-038-2

Query Match 44.2%; Score 49.5; DB 16; Length 866;
Best Local Similarity 52.2%; Pred. No. 9.8;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGAATGGLLYFRAN-GVQYKG 22
DB 617 TGLRTIDGKKYFNTNTAVATG 639

RESULT 6

US-10-011-366-6

Sequence 6, Application US/10011366

Publication No. US20030054493A1

GENERAL INFORMATION:

APPLICANT: Williams, James A.

Kink, John A.

TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES

OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE

DISEASE

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESS: Medlen & Carroll

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/011,366

FILING DATE: 16-NO. US20030054493A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/957,310

FILING DATE: 23-OCT-1997

APPLICATION NUMBER: US 08/329,154

FILING DATE: 24-OCT-1994

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPBD-01121

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-011-366-6

Query Match 44.2%; Score 49.5; DB 15; Length 2710;
Best Local Similarity 52.2%; Pred. No. 38;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGAATGGLLYFRAN-GVQYKG 22
DB 2460 TGLRTIDGKKYFNTNTAVATG 2482

RESULT 7

US-10-275-360-13

Sequence 13, Application US/10275360

Publication No. US20030115634A1

GENERAL INFORMATION:

APPLICANT: JOMAA, Haasaa

TITLE OF INVENTION: GENES OF THE 1-DESOXY-D-XYULOSE BIOSYNTHESIS PATH

FILE REFERENCE: JOMAA-12 (PCT)

CURRENT APPLICATION NUMBER: US/10/275,360

CURRENT FILING DATE: 2002-11-05

PRIOR APPLICATION NUMBER: PCT/EP01/04537

PRIOR FILING DATE: 2001-04-21

PRIOR APPLICATION NUMBER: DE10021698.9

PRIOR FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn version 3.1

SEQ ID NO 13

LENGTH: 343

TYPE: PRT

ORGANISM: Plasmodium falciparum

US-10-275-360-13

Query Match 43.8%; Score 49; DB 15; Length 343;
Best Local Similarity 56.2%; Pred. No. 3.9;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 INGOLLYFRANGVQVK 21
DB 138 ITDQLLYFGSKGVYDIK 153

RESULT 8

US-10-156-761-13796

Sequence 13796, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 13796
LENGTH: 120
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-13796

Query Match 42.9%; Score 48; DB 15; Length 120;
Best Local Similarity 40.9%; Pred. No. 1.7;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
DB 32 SGAPVYTGKVLMSADGKQVRG 53

RESULT 9
US-09-882-774-1
Sequence 1, Application US/09882774
Publication No. US20030021795A1
GENERAL INFORMATION:
APPLICANT: Houston, Michael E.
TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
FILE REFERENCE: 003592-007
CURRENT APPLICATION NUMBER: US/09/882,774
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,892
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/213,387
PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 619
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match 41.1%; Score 46; DB 11; Length 619;
Best Local Similarity 40.9%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
DB 462 TGMAYNGSWYIYNANGAMATG 483

RESULT 10
US-09-765-272-160
Sequence 160, Application US/09765272
Patent No. US20020061545A1
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland

COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 641 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-765-272-160

Query Match 41.1%; Score 46; DB 9; Length 641;
Best Local Similarity 40.9%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
DB 544 TGMAYNGSWYIYNANGAMATG 565

RESULT 11
US-09-769-787-184
Sequence 184, Application US/09769787
Publication No. US20030091577A1
GENERAL INFORMATION:
APPLICANT: Microbial Technics Limited
APPLICANT: Gilbert, Christophe FG
APPLICANT: Hansbro, Philip M
TITLE OF INVENTION: Proteins
FILE REFERENCE: PWC/P21129MO
CURRENT APPLICATION NUMBER: US/09/769,787
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: GB 9816337.1
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: US 60/125164
PRIOR FILING DATE: 1998-03-19
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 184
LENGTH: 744
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-769-787-184

Query Match 41.1%; Score 46; DB 11; Length 744;
Best Local Similarity 40.9%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
DB 647 TGMAYNGSWYIYNANGAMATG 668

```

RESULT 12
US-09-764-869-970
; Sequence 970, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 970
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-970

Query Match          39.3%; Score 44; DB 9; Length 115;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 TINGOLLYER 14
DB 63 TVNGELFLFR 72

RESULT 13
US-10-091-504-970
; Sequence 970, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1

```

```

; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 970
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-504-970

Query Match          39.3%; Score 44; DB 15; Length 115;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 TINGOLLYER 14
DB 63 TVNGELFLFR 72

RESULT 14
US-09-738-626-4344
; Sequence 4344, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYPEPTIDES

```

```

; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentln ver. 3.0
; SEQ ID NO 4344
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4344

```

```

Query Match      38.8%; Score 43.5; DB 10; Length 196;
Best Local Similarity 54.5%; Pred. No.18;
Matches 12; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

```

```

QY      1 TGARTINGQLLYFRANGVQVKG 22
      ||| ||| ||| ||| ||| |||
Db      169 TGA-TISASANTVLRAGVQVKG 189

```

```

RESULT 15
US-09-287-070-2
; Sequence 2, Application US/09287070A
; Patent No. US20020041881A1
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Gosink, Khosheh
; APPLICANT: Masure, Robert
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF NOVEL
; TITLE OF INVENTION: PNEUMOCOCCAL CHOLINE BINDING PROTERINS CbpG, AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USES THEREOF
; FILE REFERENCE: 1340-1-024cid
; CURRENT APPLICATION NUMBER: US/09/287,070A
; CURRENT FILING DATE: 1999-04-06
; EARLIER APPLICATION NUMBER: 09/196,389
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pneumococcus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (35)
; OTHER INFORMATION: The amino acid can be any of Ile, Thr, Asn, or
; OTHER INFORMATION: Ser. The DNA coding is AsT for this amino acid
US-09-287-070-2

```

```

Query Match      38.4%; Score 43; DB 9; Length 153;
Best Local Similarity 36.4%; Pred. No.16;
Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

```

```

QY      1 TGARTINGQLLYFRANGVQVKG 22
      ||| ||| ||| ||| ||| |||
Db      114 TDMQKVGKMYLNSNGAMVTG 135

```

Search completed: November 13, 2003, 10:29:01
 Job time : 29.7773 secs

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 SQ Sequence 1017 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 3.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHLSILEAWSNDNTPYLHD 21
 DB 447 ANDHLSILEAWSNDNTPYLHD 467
 RESULT 2
 AAU98027 standard; Protein; 1475 AA.
 AC AAU98027;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 DR N-PSDB; ABR52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 CC The invention of an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589B, N471D, N471D/T589D, and
 CC N471D/T589B. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising the mutant GTF, wild type or, starch, a latex,
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 CC
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHLSILEAWSNDNTPYLHD 21
 DB 481 ANDHLSILEAWSNDNTPYLHD 501
 RESULT 3
 AAU98030 standard; Protein; 1475 AA.
 AC AAU98030;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB mutant I448V.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 XX
 PI Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 DR
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98C27
 CC and the information in claim 36.
 XX
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Beet.Local Similarity 100.0%; Pred.No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHLSTLEAWSDNDPTLYLHD 21
 DB 481 ANDHLSTLEAWSDNDPTLYLHD 501
 RESULT 4
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX
 DT 27-AUG-2002 (first entry)

XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 XX
 XX
 XX US2002031826-A1.
 XX
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTYLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTYLHD 501

RESULT 5
 AAU98032 standard; Protein; 1475 AA.

AC AAU98032;
 DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T.

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

XX US2002031826-A1.

XX 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

DR WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T,
 CC Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein of potato or the vacuole of sugar beet.

CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

SO Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTYLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTYLHD 501

RESULT 6
 AAU98033 standard; Protein; 1475 AA.

AC AAU98033;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant K1014T.

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"

XX US2002031826-A1.

XX 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI; 2002-414332/44.
 DR
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -: 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, 1589E, N471D, N471D/1589D, and
 CC N471D/1589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHSLIEAMSDNDPTPLHD 21
 DB 481 ANDHSLIEAMSDNDPTPLHD 501
 RESULT 7
 AAU98034
 ID AAU98034 standard; Protein; 1475 AA.
 XX
 AC AAU98034;

XX
 DT 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N/D567T.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 457 /note= "wild-type Asp substituted by Asn"
 FT /note= "wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "wild-type Asp substituted by Thr"
 FT
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -: 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, 1589E, N471D, N471D/1589D, and
 CC N471D/1589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilising the glucan produced by GTF, which utilises
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHSLTEAMSDNDPTPLHD 21
 DB 481 ANDHSLTEAMSDNDPTPLHD 501

RESULT 8
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.

XX AC AAU98035;

XX DT 27-AUG-2002 (first entry)

XX DE S. mutans glucosyltransferase GTFB mutant D457N/D571K.

XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;

XX KW amyloplast; vacuole; paper manufacture; mutant; mutain.

XX OS Streptococcus mutans.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FT US2002031826-A1.

XX PD 14-MAR-2002.

XX PF 19-DEC-2000; 2000US-0740274.

XX PR 11-DEC-1998; 98US-0210361.

XX PR 07-JUN-1995; 95US-0478704.

XX PR 07-JUN-1995; 95US-0482711.

XX PR 07-JUN-1995; 95US-0485243.

XX PR 16-JAN-1998; 98US-0007999.

XX PR 16-JAN-1998; 98US-0008172.

XX PR 20-JUN-1998; 98US-0009620.

XX PA (NICH/) NICHOLS S E.

XX PI Nichols SE;

XX DR WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -

XX PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779C/K1014T,
 CC Y169A/Y170A/Y171A, and K779C or a GTF D polypeptide having and
 CC changes at positions from T589E, T589E, N471D, N471D/T589E and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerative plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilising the glucan produced by GTF, which utilises
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHSLTEAMSDNDPTPLHD 21
 DB 481 ANDHSLTEAMSDNDPTPLHD 501

RESULT 9

AAU98036
 ID AAU98036 standard; Protein; 1475 AA.

XX AC AAU98036;

XX DT 27-AUG-2002 (first entry)

XX DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.

XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;

XX KW amyloplast; vacuole; paper manufacture; mutant; mutain.

XX OS Streptococcus mutans.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FT US2002031826-A1.

14-MAR-2002.
19-DEC-2000; 2000US-0740274.
11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.
07-JUN-1995; 95US-0485243.
16-JAN-1998; 98US-0007999.
16-JAN-1998; 98US-0008172.
20-JAN-1998; 98US-0009620.
(NICH/) NICHOLS S E.
Nichols SE;
WPI; 2002-414332/44.
Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
Claim 36; Page -; 44pp; English.
The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations and/or vacuole or a maize line the glucan is produced in the amyloplast and/or vacuole of a plant deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;
Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ANDHLSLEAMSDNDPPIAHD 21
|||||

481 ANDHLSLEAMSDNDPPIAHD 501
RESULT 10
AAU98037
ID AAU98037 standard; Protein; 1475 AA.
XX
AC AAU98037;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
XX
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; muten.
XX
OS Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
FT
XX US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX
PI Nichols SE;
XX
DR WPI; 2002-414332/44.
XX
PT Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
XX
PS Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations and/or vacuole of a maize line the glucan is produced in the amyloplast and/or vacuole of a plant deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX Sequence 1475 AA;
 SQ
 Qy Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ANDHSLIEAWSDNDTPYLHD 21
 481 ANDHSLIEAWSDNDTPYLHD 501
 Db
 RESULT 11
 AAU98038
 ID AAU98038 standard; Protein; 1475 AA.
 AC AAU98038;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans GTFB mutant 1448V/D457N/D567T/D571K/K779G/K1014T.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutcin.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX PD 14-MAR-2002.
 XX PF 19-DEC-2000; 2000US-0740274.
 XX PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779G/K1014T,
 CC Y169N/Y170A/Y171A, and K779G or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX Sequence 1475 AA;
 SQ
 Qy Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ANDHSLIEAWSDNDTPYLHD 21
 481 ANDHSLIEAWSDNDTPYLHD 501
 Db
 RESULT 12
 AAU98039
 ID AAU98039 standard; Protein; 1475 AA.
 AC AAU98039;
 XX
 XX

27-AUG-2002 (first entry)
 S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.
 Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 coating composition; glucan; starch; latex; thermoplastic molecule;
 amyloplast; vacuole; paper manufacture; mutant; mutein.
 Streptococcus mutans.
 Synthetic.
 Key Location/Qualifiers
 Misc-difference 169..171
 /note="Wild-type Tyr-Tyr-Tyr substituted by
 Ala-Ala-Ala"
 US2002031826-A1.
 14-MAR-2002.
 19-DEC-2000; 2000US-0740274.
 11-DEC-1998; 98US-0210361.
 07-JUN-1995; 95US-0478704.
 07-JUN-1995; 95US-0482711.
 07-JUN-1995; 95US-0485243.
 16-JAN-1998; 98US-0007999.
 16-JAN-1998; 98US-0008172.
 20-JAN-1998; 98US-0009620.
 (NICH/) NICHOLS S E.
 Nichols SE;
 WPI; 2002-414332/44.
 Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 Claim 36; Page -; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K799Q/K1014T,
 Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 changes at positions from T589D, T589E, N471D, N471D/T589D, and
 N471D/T589E. Also included are a glucan produced by the GTF mutant,
 an isolated polynucleotide which encodes P1 or P2, or its complementary
 polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 an expression cassette comprising the polynucleotide operably linked to a
 promoter, a vector comprising the expression cassette, host cell
 introduced with the vector, a transgenic plant comprising the
 vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 coating composition comprising a glucan produced in a plant transformed
 with a gene encoding the mutant GTF, wild type or, starch, a latex,
 thermoplastic molecule or their combinations or glucan and starch where
 the glucan is produced in the amyloplast and/or vacuole or a maize line
 deficient in starch biosynthesis, transformed with a gene encoding a
 glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 comprising the glucan (paper sizing/coating agent). The vector is useful
 for producing a glucan in a plant. The method comprises transforming a
 plant cell with the vector, growing the plant cell under plant growing
 conditions to produce a regenerated plant and inducing expression of the
 polynucleotide for a time sufficient to produce the glucan in the
 regenerated plant, where the vector contains a transit sequence from
 ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 chlorophyll AB binding protein to produce a transgenic plant, and glucan
 is produced in the amyloplast of potato or the vacuole of sugar beet.
 Glucans are useful as substitutes for and additions to modified starch
 and latexes in paper manufacture. Unlike prior art techniques, which
 require input materials that produce chemical effluents, paper

Manufacture utilizing the glucan produced by GTF, which utilizes
 biologically produced input materials, is more cost-effective and
 environmentally friendly. Moreover, glucans also exhibit thermoplastic
 properties and impart gloss to the paper during coating step.
 The present sequence represents a GTFB mutant of the invention.
 Note: The present sequence is not shown in the specification but
 was created by the indexer using the GTFB sequence appearing as AAU98027
 and the information in claim 36.
 Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHLSILEAWSDNDPTPLHD 21
 DB 481 ANDHLSILEAWSDNDPTPLHD 501
 RESULT 13
 AAU98040
 ID AAU98040 standard; Protein; 1475 AA.
 AC AAU98040;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB mutant K779Q.
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 coating composition; glucan; starch; latex; thermoplastic molecule;
 amyloplast; vacuole; paper manufacture; mutant; mutein.
 Streptococcus mutans.
 Synthetic.
 Key Location/Qualifiers
 Misc-difference 779
 /note="Wild-type Lys substituted by Gln"
 US2002031826-A1.
 14-MAR-2002.
 19-DEC-2000; 2000US-0740274.
 11-DEC-1998; 98US-0210361.
 07-JUN-1995; 95US-0478704.
 07-JUN-1995; 95US-0482711.
 07-JUN-1995; 95US-0485243.
 16-JAN-1998; 98US-0007999.
 16-JAN-1998; 98US-0008172.
 20-JAN-1998; 98US-0009620.
 (NICH/) NICHOLS S E.
 Nichols SE;
 WPI; 2002-414332/44.
 Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 Claim 36; Page -; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K799Q/K1014T,
 Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations and/or vacuole or a maize line
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSDNDTPYLHD 21
 DB 481 ANDHSLILEAWSDNDTPYLHD 501

RESULT 14

AAU79284
 ID AAU79284 standard; Protein; 1476 AA.

AC AAU79284;

DT 13-AUG-2002 (first entry)

DE Streptococcus mutans monoclonal antibody-related protein #1.

KW Antibody; dental caries; water insoluble glucan synthetase;

KW anti-carries; glucosyl transferase-B; immunotherapy.

OS Streptococcus mutans.

PN JP2002114709-A.

PD 16-APR-2002.

PF 04-OCT-2000; 2000UP-0304889.

PR 04-OCT-2000; 2000UP-0304889.

PA (UYNI-) UNIV NIPPON.

XX WPI; 2002-448101/48.

DR WPI; 2002-448101/48.

PT Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 PS Claim 3; Page 13-16; 28pp; Japanese.

CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.

CC Sequence 1476 AA;

Query Match 100.0%; Score 119; DB 23; Length 1476;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSDNDTPYLHD 21
 DB 481 ANDHSLILEAWSDNDTPYLHD 501

RESULT 15

AAU98028
 ID AAU98028 standard; Protein; 1375 AA.

AC AAU98028;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB.

KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KW amyloplast; vacuole; paper manufacture.

OS Streptococcus mutans.

PN US2002031826-A1.

PD 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

PA (NICH/) NICHOLS S E.

PI Nichols SE;

DR WPI; 2002-414332/44.

DR N-PSDB; ABR52939.

PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -

PS Disclosure; Page 30-33; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTRC.

XX Sequence 1375 AA;

Query Match 92.4%; Score 110; DB 23; Length 1375;
 Best Local Similarity 95.2%; Pred. No. 1.1e-07;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSTLEAWSDNPPYIHD 21
 |||||
 Db 507 ANDHLSTLEAWSYNDPPYIHD 527

Search completed: November 13, 2003, 09:38:26
 Job time : 38.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154,898 Million cell updates/sec

Title: US-09-290-049A-10
Perfect score: 119
Sequence: 1 ANDHSLIEAMSDNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR.76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1475	2 B33135	gtfB protein precursor
2	110	92.4	1375	2 J70345	dextranucrase (EC
3	108	90.8	1592	2 A38175	glucosyltransferase
4	78	65.5	1365	2 A41483	glucosyltransferase
5	75	63.0	1431	2 A45866	dextranucrase (EC
6	74	62.2	1508	2 T31098	probable dextran
7	66	55.5	1449	2 T30857	glucosyltransferase
8	66	55.5	1449	2 T30852	glucosyltransferase
9	63	52.9	1577	2 T30858	glucosyltransferase
10	58	48.7	1518	2 A44811	glyceroldehyde-3-P
11	57	47.9	1331	2 B48445	glucosyltransferase
12	57	47.9	1599	2 S22737	glucosyltransferase
13	52	43.7	1488	2 T30289	pristinamycin I sy
14	50	42.0	347	2 T48610	hypothetical prote
15	49.5	41.6	549	2 T03983	hypothetical prote
16	49	41.2	302	2 AG0597	conserved hypotret
17	49	41.2	336	2 OCB840	BGLF2 protein - hu
18	49	41.2	601	2 B87028	pyruvate, phosphat
19	48	40.3	490	2 H70538	probable ppk prot
20	47.5	39.9	331	1 DEUTGC	glyceroldehyde-3-P
21	47.5	39.9	519	2 C86372	hypothetical prote
22	47	39.5	175	2 C86205	hypothetical prote
23	47	39.5	378	2 T04254	hypothetical prote
24	47	39.5	524	2 D82220	conserved hypotret
25	47	39.5	525	2 T40088	RhoGEF domain cont
26	46.5	39.1	418	2 D90506	4-aminobutyrate am
27	46	38.7	335	2 S55978	hypothetical prote
28	46	38.7	335	2 E86568	glyceroldehyde-3-P
29	46	38.7	335	2 B72053	glyceroldehyde-3-P

30	46	38.7	400	2 F98138	hypothetical prote
31	46	38.7	417	2 A83149	glucosyltransferase
32	46	38.7	476	1 S58229	salicylate biosynt
33	45.5	38.2	96	2 E81786	conserved hypotret
34	45	37.8	148	2 B95908	hypothetical prote
35	45	37.8	236	2 D65905	guanidinocetate N
36	45	37.8	309	2 D86362	conserved hypotret
37	45	37.8	947	2 B86362	hypothetical prote
38	45	37.8	1090	2 AG1749	glycosidase homolo
39	45	37.8	1091	2 AF1380	glycosidase homolo
40	45	37.8	1159	2 T43461	probable phosphodi
41	44.5	37.4	210	2 H83332	conserved hypotret
42	44	37.0	155	2 AD2392	transcription regu
43	44	37.0	247	2 PQ0178	glyceroldehyde-3-P
44	44	37.0	331	1 DEECG3	glyceroldehyde-3-P
45	44	37.0	331	2 H90939	glyceroldehyde-3-P

ALIGNMENTS

RESULT 1

B33135 gtfB protein precursor - Streptococcus mutans

C/Species: Streptococcus mutans C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999

C/Accession: B33135; A33128 R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.

J. Bacteriol. 169, 4263-4270, 1987

A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.

A/Reference number: A33135; PMID:87308013; PMID:3040685

A/Accession: B33135 A/Status: preliminary

A/Molecule type: DNA A/Residues: 1-1475 <SH1>

A/Cross-references: GB:M17361; NID:G153639; PIDN:AAA88588.1; PID:G153640

R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K. submitted to the Protein Sequence Database, September 1990

A/Reference number: A33128 A/Accession: A33128

A/Status: preliminary; not compared with conceptual translation A/Molecule type: DNA

A/Residues: 1-171,173-641,'N',643-1475 <SH2> A/Experimental source: strain GS-5

C/Superfamily: cpl repeat homology C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999

F:1096-1115/Domain: cpl repeat homology <CP1> F:1224-1243/Domain: cpl repeat homology <CP2>

F:1289-1308/Domain: cpl repeat homology <CP3> F:1354-1373/Domain: cpl repeat homology <CP4>

F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match Best Local Similarity 100.0%; Score 119; DB 2; Length 1475; Pred. No. 5.2e-10; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANDHSLIEAMSDNDPTPLHD 21
Db 481 ANDHSLIEAMSDNDPTPLHD 501

RESULT 2

J70345 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)

N/Alternate names: sucrose 6-glucosyltransferase C/Species: Streptococcus mutans

C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999 C/Accession: J70345; G33335

R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K. Gene 69, 101-109, 1988

A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.

A/Reference number: J70345; PMID:89137980; PMID:2976010

A/Accession: J70345 A/Molecule type: DNA

A;Residues: 1-1375 <UED>
A;Experimental source: GS-5
R;Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
U. Bacteriol. 169, 4263-4270, 1987
A;Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A;Reference number: A33135, MUID:87308013, PMID:3040685
A;Accession: C33135
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <SHI>
A;Cross-references: GB:M17361
C;Genetics:
A;Gene: gtfC
C;Function:
A;Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C;Superfamily: cpl repeat homology
C;Keywords: duplication; glycosyltransferase; hexosyltransferase
F;1-34/Domain: signal sequence #status predicted <SIG>
F;35-1375/Product: glucosyltransferase #status predicted <MAT>
F;1126-1145/Domain: cpl repeat homology <CP1>
F;1253-1272/Domain: cpl repeat homology <CP2>
F;1318-1337/Domain: cpl repeat homology <CP3>

Query Match 92.4%; Score 110; DB 2; Length 1375;
Best Local Similarity 95.2%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
Db 507 ANDHSLILEAWSNDNTPYLHD 527

RESULT 3
A38175
glucosyltransferase precursor - Streptococcus sobrinus
C;Species: Streptococcus sobrinus
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1995
C;Accession: A38175
R;Abo, H.; Mesumura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
U. Bacteriol. 173, 989-996, 1991
A;Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A;Reference number: A38175, MUID:91123227, PMID:1704006
A;Accession: A38175
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1592 <ABO>
A;Cross-references: GB:D90213, NID:9217032, PIDN:BA14241.1, PID:41014946, PID:9217033
C;Superfamily: cpl repeat homology
F;1093-1112/Domain: cpl repeat homology <CP1>
F;1222-1241/Domain: cpl repeat homology <CP2>
F;1287-1306/Domain: cpl repeat homology <CP3>
F;1330-1351/Domain: cpl repeat homology <CP4>
F;1352-1371/Domain: cpl repeat homology <CP5>
F;1402-1420/Domain: cpl repeat homology <CP6>
F;1465-1484/Domain: cpl repeat homology <CP7>
F;1513-1532/Domain: cpl repeat homology <CP8>

Query Match 90.8%; Score 108; DB 2; Length 1592;
Best Local Similarity 85.7%; Pred. No. 3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
Db 477 ANNHVSIVEAWSNDNTPYLHD 497

RESULT 4
A41483
glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus
C;Species: Streptococcus sobrinus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C;Accession: A41483
R;Gillmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

Infect. Immun. 58, 2452-2458, 1990
A;Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltrans
A;Reference number: A41483, MUID:90316655, PMID:2142479
A;Accession: A41483
A;Molecule type: DNA
A;Residues: 1-1365 <GIL>
A;Cross-references: GB:M30943, NID:9153652, PIDN:AAA26898.1, PID:9153653
C;Genetics:
A;Gene: gtfS
C;Superfamily: cpl repeat homology
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 65.5%; Score 78; DB 2; Length 1365;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
Db 467 AIDHSLILEAWSNDNTPYLHD 487

RESULT 5
A45866
dextransucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C;Species: Streptococcus mutans
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: A45866
R;Honda, O.; Kato, C.; Kuramitsu, H.K.
U. Gen. Microbiol. 136, 2099-2105, 1990
A;Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
A;Reference number: A45866, MUID:9110058, PMID:2148600
A;Accession: A45866
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1431 <HON>
A;Cross-references: GB:M29296
C;Superfamily: cpl repeat homology
C;Keywords: glycosyltransferase; hexosyltransferase
F;181-201/Domain: cpl repeat homology <CP1>
F;1127-1146/Domain: cpl repeat homology <CP2>
F;1192-1211/Domain: cpl repeat homology <CP3>
F;1257-1276/Domain: cpl repeat homology <CP4>
F;1277-1297/Domain: cpl repeat homology <CP5>
F;1321-1340/Domain: cpl repeat homology <CP6>
F;1341-1361/Domain: cpl repeat homology <CP7>
F;1385-1404/Domain: cpl repeat homology <CP8>

Query Match 63.0%; Score 75; DB 2; Length 1431;
Best Local Similarity 71.4%; Pred. No. 0.0036;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
Db 495 AINHLILEAWSNDNTPYLHD 515

RESULT 6
T31098
probable dextransucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
C;Species: Leuconostoc mesenteroides
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C;Accession: T31098
R;Monchois, V.; Remaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A;Title: Cloning and sequencing of a gene coding for an extracellular dextransucrase (D
A;Reference number: Z20981, MUID:98164374, PMID:9503626
A;Accession: T31098
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: DNA
A;Residues: 1-1508 <MON>
A;Cross-references: EMBL:AF030129, NID:92766611, PID:92766612, PIDN:AAB95453.1
A;Experimental source: strain NRRL B-1299
C;Genetics:

A:Gene: darB
C:Function:
A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.0055;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLLEAMSDNDTPYLHD 21
DB 563 ANQHSLLEAMSDNDPEYKCD 583

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30857
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-1449 <SIM>
A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AA041412.1
C:Genetics:
A:Gene: gtfI

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPYLHD 21
DB 539 HSLLEAMSHNDAYNED 556

RESULT 8
T30552

glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaife, R.I.

A:Description: The EMBL Data Library, February 1998
submitted to the EMBL Data Library, February 1998.
A:Reference number: Z20854

A:Accession: T30552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <GAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AA05156.1
C:Genetics:
A:Gene: gtfN

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPYLHD 21
DB 539 HSLLEAMSHNDAYNED 556

RESULT 9
T30858

glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>
A:Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIDN:AA041413.1
C:Genetics:
A:Gene: gtfm

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.3;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPYLHD 21
DB 591 HSLLEAMSYNDHQYKCD 608

RESULT 10
A44811

glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809

R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811

A:Molecule type: DNA
A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:211873; NID:947526; PIDN:CAA77900.1; PID:g47527
A:Note: sequence extracted from NCBI backbone (NCBIN:81050, NCBI:P:81052)

C:Genetics:
A:Gene: gtfJ
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 48.7%; Score 58; DB 2; Length 1518;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPY 18
DB 504 HSLVLEAMSLNDNH 518

RESULT 11
B48445

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me
C:Species: Leishmania mexicana
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: B48445; S25142

R:Hanaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Oppendes, F.R.; Michels, P.A.M.
Mol. Biochem. Parasitol. 55, 115-126, 1992
A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate de

A:Reference number: A48445; MUID:93063042; PMID:1435864
A:Accession: B48445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <HAN>
A:Cross-references: EMBL:X65220; NID:g95552; PIDN:CAA46323.1; PID:g95553
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 47.9%; Score 57; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDTPYLH 20
 Db 301 NDHFKLVSWYDNETGYSH 319

RESULT 12

S22737
 glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
 C:Species: Streptococcus salivarius
 C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
 C/Accession: S22737; S28810; B44811; S22727
 R/Jacques, N.
 submitted to the EMBL Data Library, March 1992
 A/Reference number: S22726
 A/Accession: S22737
 A/Molecule type: DNA
 A/Residues: 1-1599 <JAC>
 A/Cross-references: EMBL:Z11872; NID:947530; PIDN:CAA77898.1; PID:947531
 A/Experimental source: ATCC 25975
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
 A/Reference number: A44811; MUID:92148377; PMID:1838391
 A/Accession: S28810
 A/Molecule type: DNA
 A/Residues: 1-51 <GIF>
 A/Cross-references: EMBL:Z11873
 C/Genetics:
 A/Gene: gtfK
 A/Superfamily: cpl repeat homology
 C/Keywords: glycosyltransferase; hexosyltransferase
 F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 47.9%; Score 57; DB 2; Length 1599;
 Best Local Similarity 55.6%; Pred. No. 2.6;
 Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLILEAMSDNDTPYLH 21
 Db 494 NLSILEAMSHNDPPYYNE 511

RESULT 13

T30289
 pristinaamycin I synthase 3 - Streptomyces pristinaespiralis
 C:Species: Streptomyces pristinaespiralis
 C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Dec-2000
 C/Accession: T30289
 R/de Crey-Lagard, V.A.; Saurin, W.; Thibaut, D.; Gil, P.; Naudin, L.; Crouzet, J.; Bian
 submitted to the EMBL Data Library, February 1997
 A/Description: Streptogramin B biosynthesis in Streptomyces pristinaespiralis and Strept
 A/Reference number: Z20808
 A/Accession: T30289
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-4848 <DEC>
 A/Cross-references: EMBL:Y11548; NID:e1025755; PID:e3075539; PIDN:CAA72312.1
 C/Genetics:
 A/Gene: snbDE
 C/Superfamily: acyl carrier protein homology; acetate-CoA ligase homology
 C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
 F:511-951/Domain: acetate-CoA ligase homology <ACLI1>
 F:966-1034/Domain: acyl carrier protein homology <ACRP1>
 F:1563-2024/Domain: acetate-CoA ligase homology #status atypical <ACLI2>
 F:2399-2467/Domain: acyl carrier protein homology <ACRP2>
 F:2995-3427/Domain: acetate-CoA ligase homology <ACLI3>
 F:3441-3509/Domain: acyl carrier protein homology <ACRP3>
 F:4403-4492/Domain: acetate-CoA ligase homology <ACLI4>
 F:4507-4575/Domain: acyl carrier protein homology <ACRP4>
 F:3998,2431,3473/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 43.7%; Score 52; DB 2; Length 4848;

Best Local Similarity 47.1%; Pred. No. 58;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTP 17
 Db 3972 ADDHARVLGTGMNDTDPH 3988

RESULT 14

T48610
 hypothetical protein F18022.160 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C/Accession: T48610
 R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A/Reference number: Z24493
 A/Accession: T48610
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-347 <BRV>
 A/Cross-references: EMBL:DB
 A/Experimental source: cultivar Columbia; BAC clone F18022
 C/Genetics:
 A/Map position: 5
 A/Intons: 232/3; 296/3
 A/Note: F18022.160

Query Match 42.0%; Score 50; DB 2; Length 347;
 Best Local Similarity 52.9%; Pred. No. 5.2;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 DHTSILEAMSDNDTPYL 19
 Db 200 DYEQIMEAMSDKGTLVY 216

RESULT 15

T03983
 r22 nuclear reatorer protein - maize
 N/Alternate names: aldehyde dehydrogenase homolog
 C/Species: Zea mays (maize)
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C/Accession: T03983
 R/Cui, X.; Wise, R.P.; Schnable, P.S.
 Science 272, 1334-1336, 1996
 A/Title: The r22 nuclear reatorer gene of male-sterile T-cytoplasm maize.
 A/Reference number: Z15177; MUID:96243131; PMID:8650543
 A/Accession: T03983
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-549 <CUV>
 A/Cross-references: EMBL:U43082; NID:g1421729; PIDN:AAC49371.1; PID:g1421730
 A/Experimental source: strain B73
 C/Genetics:
 A/Map position: 9
 A/Note: r22
 C/Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
 F:107-370/Domain: aldehyde dehydrogenase homology <ALD>

Query Match 41.6%; Score 49.5; DB 2; Length 549;
 Best Local Similarity 58.8%; Pred. No. 11;
 Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHLSILEAMSDNDTPY 18
 Db 147 NDELAALETW-DNGKPY 162

Search completed: November 13, 2003, 09:50:24
 Job time : 13.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049A-10

Perfect score: 119
Sequence: 1 ANDHUSITLFAWSDNDTPYLDH 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1476	1	GPFB_STRMU
2	110	92.4	1455	1	GPFC_STRMU
3	108	90.8	1592	1	GPFI_STRMU
4	108	90.8	1597	1	GPFI_STRMU
5	78	65.5	1365	1	GPFS_STRMU
6	75	63.0	1462	1	GPFD_STRMU
7	57	47.9	1330	1	GPFC_STRMU
8	49	41.2	302	1	YBHK_SALTY
9	49	41.2	302	1	YBHK_SALTY
10	49	41.2	336	1	U1A6_EBV
11	47.5	39.9	330	1	GPFC_STRMU
12	47	39.5	337	1	GPFC_STRMU
13	46	38.7	275	1	SC35_YEAST
14	46	38.7	335	1	GPFC_STRMU
15	46	38.7	476	1	PCNA_PSEAR
16	45	37.8	236	1	GPFC_STRMU
17	45	37.8	385	1	GPFC_STRMU
18	44.5	37.4	470	1	AMP2_STRMU
19	44	37.0	234	1	GPFC_STRMU
20	44	37.0	330	1	GPFC_STRMU
21	44	37.0	330	1	GPFC_STRMU
22	44	37.0	332	1	GPFC_STRMU
23	44	37.0	332	1	GPFC_STRMU
24	44	37.0	337	1	GPFC_STRMU
25	44	37.0	337	1	GPFC_STRMU
26	44	37.0	337	1	GPFC_STRMU
27	44	37.0	337	1	GPFC_STRMU
28	44	37.0	337	1	GPFC_STRMU
29	44	37.0	337	1	GPFC_STRMU
30	43.5	36.6	115	1	TKN1_RABIT
31	43.5	36.6	130	1	TKN1_MESAT
32	43.5	36.6	130	1	TKN1_MOUSE
33	43.5	36.6	130	1	TKN1_MOUSE

34	43.5	36.6	491	1	GPFD_ERWCH
35	43	36.1	174	1	PADC_VIBCH
36	43	36.1	235	1	GPFC_STRMU
37	43	36.1	236	1	GPFC_STRMU
38	43	36.1	320	1	GPFC_STRMU
39	43	36.1	322	1	GPFC_STRMU
40	43	36.1	332	1	GPFC_STRMU
41	43	36.1	332	1	GPFC_STRMU
42	43	36.1	334	1	GPFC_STRMU
43	43	36.1	337	1	GPFC_STRMU
44	43	36.1	337	1	GPFC_STRMU
45	43	36.1	424	1	IVD_RAT

ALIGNMENTS

RESULT 1
ID GPFB_STRMU STANDARD: PRT; 1476 AA.
AC P08987; 069381; 069384; 069387; 069390; 069396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
GN (Sucrose 6-glucosyltransferase).
OS GPFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5.
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans."
RL J. Bacteriol. 169:4263-4270(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype B, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL FEWS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Aylce D., McNamee W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Jia H., Jia Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Fretwell J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
fructose + (1,6)-alpha-D-glucosyl (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation. Its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M17361; AAA89586.1; -
CC DR EMBL; D88651; BAA26101.1; -
CC DR EMBL; D88654; BAA26105.1; -
CC DR EMBL; D88657; BAA26109.1; -
CC DR EMBL; D88660; BAA26113.1; -
CC DR EMBL; D89977; BAA26119.1; -
CC DR EMBL; AE014940; AAN59705.1; -
CC DR InterPro; IPR002479; CM_binding_70.
CC DR InterPro; IPR003318; Glyco_hydro_70.
CC DR Pfam; PF02324; Glyco_hydro_70; 1.
CC DR Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
CC Complete proteome.
CC KM
CC FT SIGNAL 1 34
CC FT CHAIN 35 1476
CC FT DOMAIN 35 1051
CC FT DOMAIN 1097 1476
CC FT REPEAT 1097 1130
CC FT DOMAIN 1161 1470
CC FT REPEAT 1161 1210
CC FT REPEAT 1161 1275
CC FT REPEAT 1290 1340
CC FT REPEAT 1355 1405
CC FT REPEAT 1420 1470
CC FT REPEAT 62 62
CC FT VARIANT 65 65
CC FT VARIANT 68 68
CC FT VARIANT 78 78
CC FT VARIANT 86 86
CC FT VARIANT 89 89
CC FT VARIANT 168 168
CC FT VARIANT 276 276
CC FT VARIANT 399 399
CC FT VARIANT 474 474
CC FT VARIANT 512 512
CC FT VARIANT 519 519
CC FT VARIANT 701 701
CC FT VARIANT 708 708
CC FT VARIANT 938 938
CC FT VARIANT 952 957
CC FT VARIANT 963 964
CC FT VARIANT 968 970
CC FT VARIANT 1086 1086
CC FT VARIANT 1158 1158
CC FT VARIANT 1163 1163
CC FT VARIANT 1168 1168
CC FT VARIANT 1182 1182
CC FT VARIANT 1234 1234
CC FT VARIANT 1263 1263
CC FT VARIANT 1263 1263
CC FT VARIANT 1264 1264
CC FT VARIANT 1272 1272
CC FT VARIANT 1329 1329
CC FT VARIANT 1394 1394
CC FT VARIANT 1402 1402
CC -----
CC POTENTIAL.
CC GLUCOSYLTRANSFERASE-L.
CC CATALYTIC (APPROXIMATE).
CC GLUCAN-BINDING (APPROXIMATE).
CC A REPEAT.
CC 5 X TANDEM REPEATS.
CC 1.
CC 2.
CC 3.
CC 4.
CC 5.
CC S -> T (IN STRAIN MT4239).
CC T -> I (IN STRAIN GS-5).
CC V -> A (IN STRAINS GS-5, MT4245, MT4251,
CC MT4467 AND MT8148).
CC Q -> P (IN STRAIN MT4251).
CC I -> S (IN STRAINS GS-5, MT4245, MT4251,
CC MT4467 AND MT8148).
CC K -> F (IN STRAIN MT4251).
CC S -> N (IN STRAIN MT4251).
CC S -> D (IN STRAINS GS-5, MT4467 AND
CC MT8148).
CC N -> R (IN STRAIN MT4239).
CC I -> T (IN STRAIN MT4239).
CC K -> R (IN STRAIN MT8148).
CC F -> Y (IN STRAIN MT8148).
CC T -> I (IN STRAIN MT8148).
CC A -> V (IN STRAIN MT8148).
CC F -> L (IN STRAIN MT8148).
CC FGKPEVE -> YGPIPIVA (IN STRAINS GS-5, MT4239
CC AND MT4467).
CC SV -> NT (IN STRAINS GS-5, MT4239 AND
CC MT4467).
CC ADS -> VDG (IN STRAINS GS-5, MT4239 AND
CC MT4467).
CC A -> T (IN STRAIN MT4239).
CC S -> N (IN STRAIN MT4239).
CC H -> Y (IN STRAIN MT4251).
CC H -> K (IN STRAIN MT8148).
CC Y -> C (IN STRAIN MT8148).
CC R -> P (IN STRAIN MT4239).
CC R -> P (IN STRAIN MT8148).
CC R -> H (IN STRAINS GS-5 AND MT4467).
CC Y -> H (IN STRAINS GS-5, MT4239, MT4467
CC AND MT8148).
CC S -> G (IN STRAINS GS-5, MT4239, MT4467
CC AND MT8148).
CC H -> Y (IN STRAINS GS-5 AND MT4467).
CC Y -> H (IN STRAINS GS-5, MT4239, MT4467
CC AND MT8148).
CC S -> G (IN STRAINS GS-5, MT4239, MT4467

```

```

FT          VARIANT      1459      1459      Y -> H (IN STRAIN MT4467) .
FT          CONFLICT     570      570      R -> A (IN REF. 1) .
FT          CONFLICT     800      817      ADDQVRAAASPTSDTK -> LKMFALRLARPHQMA
FT                                     (IN REF. 1) .
FT          CONFLICT     1310     1310     H -> L (IN REF. 1) .
SQ          SEQUENCE     1476 AA; 165846 MW; 9CE809F731B4CBCE CRC64;

Query Match      100.0%; Score 119; DB 1; Length 1476;
Best Local Similarity 100.0%; Prid. No. 1.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 ANDHSLTEAWSDNDPTFLHD 21
DB      481 ANDHSLTEAWSDNDPTFLHD 501

RESULT 2
GTFIC_STRMU      STANDARD; PRT; 1455 AA.
ID      GTFIC_STRMU      01-NOV-1988 (Rel. 09, Created) P05427;
AC      P13470; 068382; 069385; 069388; 069391; 069397;
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE      (Debrunucrase) (Sucrose 6-glucosyltransferase) .
GN      GTFIC OR SMU.1005.
OS      Streptococcus mutans.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_TaxId=1309;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=GS-5;
RC      MEDLINE=89137980; PubMed=2976010;
RA      Ueda S., Shiroza T., Kuramitsu H.K.;
RT      "Sequence analysis of the gtfIC gene from Streptococcus mutans GS-5." ;
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MT439 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC      MT4467 / Serotype E, and MT8148 / Serotype C;
RX      MEDLINE=98231643; PubMed=9570124;
RA      Fujiwara T., Terao Y., Hoshino T., Kawabata S., Sobue S.,
RA      Kimura S., Hamada S.;
RT      "Molecular analysis of glucosyltransferase genes among strains of
RL      Streptococcus mutans." ;
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=UA159 / ATCC 700610 / Serotype C;
RX      MEDLINE=22295063; PubMed=12397186;
RA      Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA      Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA      Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perretti J.J.;
RT      "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT      pathogen." ;
RN      [4]
RP      SEQUENCE OF 1-349 FROM N.A.
RC      STRAIN=GS-5;
RC      MEDLINE=87308013; PubMed=3040685;
RA      Shiroza T., Ueda S., Kuramitsu H.K.;
RT      "Sequence analysis of the gtfB gene from Streptococcus mutans." ;
RN      [5]
RP      FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT
CC      TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC      OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC      AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC      -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
CC      fructose + [(1,6)-alpha-D-glucosyl] (N+1) .
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- DISEASE: DENTAL CARIES.

```

	-I- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-IN-SOLUBLE GLYCANS (ALPHA
	1,3-LINKED GLUCOSE AND SOME 1,6-GLUCOSE). GTF-S SYNTHESIZES
	WATER-SOLUBLE GLYCANS (ALPHA 1,6-GLUCOSE). GTF-ST SYNTHESIZES BOTH
	FORMS OF GLUCANS.
	-I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
	BINDING PROTEIN FROM S. MUTANS.
	-I- SIMILARITY: Contains 5 cell wall binding repeats.

	This SWISS-PROT entry is copyright. It is produced through a collaboration
	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
	the European Bioinformatics Institute. There are no restrictions on its
	use by non-profit institutions as long as its content is in no way
	modified and this statement is not removed. Usage by and for commercial
	entities requires a license agreement (see http://www.isb-sib.ch/announce/
	or send an email to license@isb-sib.ch).

	EMBL; M2054; AAA8592.1; .
	DR EMBL; D88652; BAA26102.1; .
	DR EMBL; D88658; BAA26106.1; .
	DR EMBL; D88658; BAA26110.1; .
	DR EMBL; D88678; BAA26114.1; .
	DR EMBL; D89978; BAA26120.1; .
	DR EMBL; AE014940; AAAS68706.1; .
	DR EMBL; M17351; AAA8589.1; .
	InterPro; IPR002479; CW_binding.
	InterPro; IPR003318; Glyco_hydro_70.
	Pfam; PF02324; Glyco_hydro_70; 1.
	KX Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
	Complete proteome.
XV SIGNAL	1 34
FT CHAIN	35 1455
FT DOMAIN	35 1050
FT FT	1126 1455
FT DOMAIN	1126 1455
FT REPEAT	1126 1159
FT REPEAT	1169 1200
PT REPEAT	1227 1238
PT REPEAT	1253 1303
PT REPEAT	1318 1330
FT VARIANT	21 21
FT VARIANT	81 81
FT VARIANT	106 106
FT VARIANT	116 116
FT VARIANT	126 126
FT VARIANT	150 151
FT VARIANT	256 256
FT VARIANT	425 425
FT VARIANT	519 519
FT VARIANT	538 538
FT VARIANT	545 545
FT VARIANT	597 597
FT VARIANT	600 600
FT VARIANT	601 601
FT VARIANT	614 614
FT VARIANT	727 727
FT VARIANT	734 734
FT VARIANT	964 964
FT VARIANT	1113 1113
FT VARIANT	1118 1118
FT VARIANT	1204 1204
FT VARIANT	1208 1208
FT VARIANT	1292 1294
FT VARIANT	1305 1369
FT VARIANT	1326 1326
FT VARIANT	1331 1331
FT VARIANT	1377 1377
R ->> V (IN STRAINS GS-5 AND MT4467) .	
R ->> N (IN STRAIN MT4251) .	
R ->> K (IN STRAINS GS-5 AND MT4251) .	
R ->> F (IN STRAINS GS-5 AND MT4251) .	
N ->> Y (IN STRAINS GS-5 AND MT4251) .	
A ->> T (IN STRAINS GS-5) .	
V ->> I (IN STRAIN GS-5) .	
DGH ->> NGY (IN STRAINS GS-5, MT4467 AND	
MTB148) .	
T ->> V (IN STRAIN MTB148) .	
I ->> A (IN STRAINS GS-5, MT4239, MT4467	
AND MTB148) .	
R ->> K (IN STRAIN MTB148) .	

```

FT 1  VARIANT 1398 1398 V -> I (IN STRAIN MT8148) .
FT 2  VARIANT 1424 1424 D -> N (IN STRAIN MT4239) .
FT 3  VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148) .
FT 4  VARIANT 1444 1444 S -> P (IN STRAIN MT8148) .
FT 5  CONFLICT 1337 1455 QBLTFKSNVQAKGELITERRKRIKTYDPSNGSEVFRNRXYR
FT 6  CONFLICT 1337 1455 TSSGMVYFGNDGYALIGWVVEGRRVYFDENGVFRVASHD
FT 7  CONFLICT 1337 1455 OGNHMDYDRDPRDGRSSSAVFRFRSRNGFEFNFRF ->
FT 8  CONFLICT 1337 1455 HASITISWFRRLRESLGSVKVYSNTMLIPKKEFIYVM
FT 9  CONFLICT 1337 1455 (IN REF. 1)
SQ SEQUENCE 1455 AA, 162965 MW, 3CB455A99A4FEC86 CRC64;
Oy 1 ANDHLSTLEAMSNDTPYLHD 21
Db 507 ANDHLSTLEAMSNDTPYLHD 527
Query Match 92.4%; Score 110; DB 1; Length 1455;
Best Local Similarity 95.2%; Pred. No. 3.3e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase) .
OS Streptococcus downei (Streptococcus sobrinus) .
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus .
OX NCBI_TaxID=1317;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=170406;
RA Abo H., Matsumura T., Kodama T., Ohka H., Fukui K., Kato K.,
RA Kagawa H.,
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT hydrolase)".
RL J. Bacteriol. 173:989-996(1991) .
CC -I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -I- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1) .
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DISEASE: DENTAL CARIES.
CC -I- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-ST SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -I- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch).
CC -----
DR EMBL: D90213; BAAL14241.1; -
DR InterPro: IPR002479; CM-binding .
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CM_binding_1_13 .
DR Pfam: PF02344; Glyco_hydro_70; 1.
KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.

```


FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
 FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
 FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
 FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
 FT REPEAT 1093 1142 1.
 FT REPEAT 1158 1207 2.
 FT REPEAT 1222 1272 3.
 FT REPEAT 1287 1337 4.
 FT REPEAT 1402 1451 5.
 FT REPEAT 1514 1563 6.
 FT REPEAT 1577 1592 7. (INCOMPLETE).
 SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351BCF CRC64;
 Query Match 90.8%; Score 108; DB 1; Length 1592;
 Best Local Similarity 85.7%; Pred. No. 7.6e-09;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLHD 21
 Db 477 ANNHVSIVEAMSDNTPYLHD 497

RESULT 4
 GTF1_STRDO STANDARD; PRT; 1597 AA.
 ID GTF1_STRDO
 AC P11001;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFI.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEF28;
 RX MEDLINE=87308014; PubMed=3040686;
 RA Ferretti J.J., Galpin M.L., Russell R.R.B.;
 RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
 RT sobrinus MEF28."
 RL J. Bacteriol. 169:4271-4278(1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
 CC fructose + (1,6)-alpha-D-glucosyl (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-SI SYNTHESIZES BOTH
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M17391; AAC63063.1; -
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco Hydro_70.
 DR Pfam; PF01473; CW_binding_1; 16.

DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
 FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
 FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
 FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
 FT REPEAT 1099 1132 A REPEAT.
 FT REPEAT 1163 1213 AC REPEAT.
 FT REPEAT 1227 1277 AC REPEAT.
 FT REPEAT 1292 1342 AC REPEAT.
 FT REPEAT 1352 1399 B REPEAT.
 FT REPEAT 1406 1455 B REPEAT.
 FT REPEAT 1465 1512 B REPEAT.
 FT REPEAT 1519 1568 AC REPEAT.
 FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
 SQ SEQUENCE 1597 AA; 177080 MW; B9B6A200868798E CRC64;
 Query Match 90.8%; Score 108; DB 1; Length 1597;
 Best Local Similarity 85.7%; Pred. No. 7.6e-09;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLHD 21
 Db 483 ANNHVSIVEAMSDNTPYLHD 503

RESULT 5
 GTF5_STRDO STANDARD; PRT; 1365 AA.
 ID GTF5_STRDO
 AC P29336;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucosyltransferase-5 precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTF5.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEF28;
 RX MEDLINE=90316665; PubMed=2142479;
 RA Gilmore K.S., Russell R.R., Ferretti J.J.;
 RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a
 RT glucosyltransferase that synthesizes soluble glucans."
 RL Infect. Immun. 58:2452-2458(1990).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
 CC fructose + (1,6)-alpha-D-glucosyl (N+1).
 CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
 CC PRIMER GLUCAN UNLIKE GTF-I.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
 CC 1,6-GLUCOSE).
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M30943; AAA6898.1; -

```

DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding 1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1
FT CHAIN 37 1365 OR 37 (POTENTIAL).
FT REPEAT 157 177 CELL WALL BINDING 1.
FT REPEAT 178 197 CELL WALL BINDING 1.
FT REPEAT 198 1961 CATALYTIC (APPROXIMATE).
FT DOMAIN 1062 1082 CELL WALL BINDING 3.
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
FT SEQUENCE 1365 AA; 151590 MW; 167296BSA2EB8C476 CRC64;

Query Match 65.5%; Score 78; DB 1; Length 1365;
Best Local Similarity 71.4%; Pred. No. 0.00037;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHLSILEKMSDNDPTVYLD 21
Db 467 AIDHLSILEKMSGNDNDYVD 487

RESULT 6
GTFD_STRMU STANDARD; PRT; 1462 AA.
ID PA9331, O69383; O69386; O69392; O69398;
AC 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
the glucosyltransferase-S enzyme."
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C; MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terzo Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

```

```

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M29296; AAA26895.1; -
CC EMBL; D88653; BAA26103.1; -
CC EMBL; D88656; BAA26107.1; -
CC EMBL; D88659; BAA26111.1; -
CC EMBL; D88662; BAA26115.1; -
CC EMBL; D89979; BAA26121.1; -
CC EMBL; AE014932; AAN58619.1; -
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
CC Complete proteome.
CC KW
CC CHAIN 1 1462
CC DOMAIN 1232 1423
CC REPEAT 1232 1295
CC REPEAT 1296 1359
CC REPEAT 1360 1423
CC REPEAT 10 10
CC VARIANT 19 19
CC VARIANT 58 58
CC VARIANT 68 68
CC VARIANT 81 81
CC VARIANT 113 113
CC VARIANT 122 122
CC VARIANT 132 132
CC VARIANT 135 135
CC VARIANT 137 137
CC VARIANT 202 202
CC VARIANT 255 255
CC VARIANT 275 275
CC VARIANT 288 288
CC VARIANT 301 301
CC VARIANT 313 313
CC VARIANT 317 317
CC VARIANT 328 328
CC VARIANT 350 350
CC VARIANT 628 633
CC VARIANT 658 668
CC VARIANT 726 732
CC VARIANT 726 730
CC VARIANT 762 762
CC VARIANT 762 762

```

```

FT  VARIANT  964  964  D -> Y (IN STRAIN MT4251)
FT  VARIANT  1019 1019  E -> K (IN STRAINS MT4245 AND MT4251)
FT  VARIANT  1059 1060  LG -> R (IN STRAIN MT4251)
FT  VARIANT  1060 1060  G -> R (IN STRAIN MT4245)
FT  VARIANT  1080 1080  G -> R (IN STRAIN MT4239)
FT  VARIANT  1142 1142  S -> Q (IN STRAIN GS-5)
FT  VARIANT  1198 1198  S -> N (IN STRAIN MT4251)
FT  VARIANT  1220 1220  Y -> C (IN STRAINS MT4251 AND MT4467)
FT  VARIANT  1280 1280  F -> L (IN STRAIN MT4467)
FT  VARIANT  1282 1282  Q -> P (IN STRAIN MT4245)
FT  VARIANT  1290 1290  K -> T (IN STRAIN MT4245)
FT  VARIANT  1311 1311  N -> D (IN STRAIN MT4245)
FT  VARIANT  1403 1403  G -> D (IN STRAINS GS-5 AND MT4467)
FT  VARIANT  1405 1405  G -> R (IN STRAIN GS-5)
FT  VARIANT  1425 1425  R -> K (IN STRAIN MT4467)
FT  VARIANT  1449 1449  RYDKNSGNVYKVVTLANGRRITGIDRMGIARY -> VY
FT  CONFLICT 1428 1462  R (IN REF. 1)
SQ  SEQUENCE 1462 AA; 163387 MW; CEA4279C4D708645 CRC64;

Query Match 63.0%; Score 75; DB 1; Length 1462;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Ox  1 ANDHSTLEAMSDNDPEYLH 21
Db  495 AINHLSTLEAMSDNDPQINKD 515

RESULT 7
G3PC_LEIME STANDARD; PRT; 330 AA.
ID  G3PC_LEIME
AC  Q01558;
DT  01-APR-1993 (Rel. 25, Created)
DI  01-FEB-1994 (Rel. 28, Last sequence update)
DE  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Glyceroldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE  (GAPDH).
GN  GAPDH.
OS  Leishmania mexicana.
OC  Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX  NCBI_Taxid=5665;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=SSP. Mexicana;
RX  MEDLINE=93063042; PubMed=1435864;
RA  Hannaert V., Blaauw M., Kohl L., Allert S., Oppenides F.R.,
RA  Michels P.A.M.;
RT  "Molecular analysis of the cytosolic and glycosomal glyceroldehyde-3-
RT  phosphate dehydrogenase in Leishmania mexicana.";
RT  Mol. Biochem. Parasitol. 55:115-126(1992).
CC  -1- CATALYTIC ACTIVITY: D-glyceroldehyde 3-phosphate + phosphate +
CC  NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC  -1- PATHWAY: Second phase of glycolysis; first step.
CC  -1- SUBUNIT: Homotrimer.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: Belongs to the glyceroldehyde 3-phosphate
CC  dehydrogenase family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: X65220; CAA46323.1; -
CC  PIR: B48445; B48445.
CC  HSP: P06977; IGAD.
CC  InterPro: IPR000173; GAP_dhdtrogenase.
CC  InterPro: IPR006424; GAPDH-1.
CC  Pfam: PF00044; gpdh. 1.
CC  Pfam: PF02800; gpdh_C. 1.

```

```

DR  PRINTS; PR00078; G3PDHDSGNASE.
DR  TIGRfam; TIGR01534; GAPDH-1; 1.
DR  PROSITE; PS00071; GAPDH; 1.
KW  Glycolysis; Oxidoreductase; NAD.
FT  INIT_MET 0
FT  BINDING 148 148 BY SIMILARITY.
FT  ACT_SITE 175 175 GLYCERALDEHYDE 3-PHOSPHATE
FT  ACT_SITE 175 175 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ  SEQUENCE 330 AA; 35511 MW; ED4B6DBE8A207F1E CRC64;

Query Match 47.9%; Score 57; DB 1; Length 330;
Best Local Similarity 52.6%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Ox  2 NDHSLTEAMSDNDPEYLH 20
Db  300 NDHFLVAVSWYNDNETGYSH 318

RESULT 8
YBHK_SALTI STANDARD; PRT; 302 AA.
ID  YBHK_SALTI
AC  P58587;
DT  28-FEB-2003 (Rel. 41, Created)
DI  28-FEB-2003 (Rel. 41, Last sequence update)
DI  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Hypothetical UPF0052 protein ybhk.
DE  YBHK OR STY0835 OR T2086.
GN  Salmonella typhi.
OS  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Salmonella.
OX  NCBI_Taxid=601;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CT18;
RX  MEDLINE=21534947; PubMed=11677608;
RA  Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA  Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA  Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA  Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA  Feltham T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
RA  Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA  Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA  Whitehead S., Barrett B.G.;
RT  "Complete genome sequence of a multiple drug resistant Salmonella
RT  enterica serovar Typhi CT18.";
RT  Nature 413:848-852(2001).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Ty2 / ATCC 700931;
RX  MEDLINE=22531367; PubMed=12644504;
RA  Deng W., Lion S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA  Burland V., Kodyanski V., Schwartz D.C., Blatter F.R.;
RT  "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT  and CT18.";
RT  Bacteriol. 185:2330-2337(2003).
CC  -1- SIMILARITY: Belongs to the UPF0052 family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: AL627268; CAD05249.1; -
CC  EMBL: AE016841; AAC069704.1; -
CC  InterPro: IPR002882; UPF0052.
CC  Pfam: PF01933; UPF0052. 1.
CC  Hypothetical protein; Complete proteome.
CC  SEQUENCE 302 AA; 32656 MW; 2218471211E185C3 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 302;

```

Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSLEAMSNDTPYLHD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 9
YBHK_SALTY STANDARD; PRT; 302 AA.
AC P58586;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
GN YBHK OR STM0801.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxId=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UT2 / SCS01412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McCallum M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.; Miller W., Stoneking T., Nhan M.,
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -1- SIMILARITY: Belongs to the UPF0052 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE008733; AAL19738.1;
DR StryGene; SG7272; ybhK.
DR InterPro: IPR002882; UPF0052.
DR Pfam; PF01933; UPF0052.1
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32626 MW; 38A84708A1E185C3 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 302;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSLEAMSNDTPYLHD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 10
UL16_EBV STANDARD; PRT; 336 AA.
AC P03221;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE BGLF2 protein.
GN BGLF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxId=10377;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bahter A.T., Biggin M.D., Deininger P.L., Farrell P.U.,
RA Gibson J.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Bartell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113548; PubMed=1662696;
RA Chen W.R., Heu T.Y., Lin S.W., Chen J.Y., Yang C.S.;
RT "Cloning and characterization of cDNA clones corresponding to
RT transcripts from the BamHI G region of the Epstein-Barr virus genome
RT and expression of BGLF2.";
RL J. Gen. Virol. 72:3047-3055(1991).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
CC HSV-6 ORF11R, EBV-1 46, HCMV UL94, EBV BDLF2, AND VZV 44.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M60514; AAA45871.1;
DR EMBL: V01555; CAA24831.1;
DR EMBL: S77132; AAB21113.1;
DR PIR; C43044; Q0BE40.
DR InterPro: IPR004286; UL16_UL94.
DR Pfam: PF03044; UL16_UL94.1.
KM Late protein.
SQ SEQUENCE 336 AA; 36888 MW; 840937A16D584C CRC64;

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 2.8;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 ANDHLSLEAMSNDTP 17
DB 216 AGAHVNLKGTWDDSP 232

RESULT 11
G3PC_TRYB STANDARD; PRT; 330 AA.
AC P10097;
DT 01-MAY-1989 (Rel. 10, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (BC 1.2.1.12)
DE (GAPDH).
OS Trypanosoma brucei brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=427;
RX MEDLINE=91249838; PubMed=2040303;
RA Michels P.A.M., Marchand M., Kohl L., Allert S., Wierenga R.K.,
RA Oppendoor F.R.;
RT "The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate
RT dehydrogenase in Trypanosoma brucei have a distant evolutionary
RT relationship.";
RL Eur. J. Biochem. 198:421-428(1991).
RN [2]
RP PRELIMINARY SEQUENCE OF 1-85.
RX MEDLINE=87161817; PubMed=3830153;
RA Misset O., van Beuningen J., Lambert A.M., van der Meer R.,
RA Oppendoor F.R.;
RT "Glyceraldehyde-phosphate dehydrogenase from Trypanosoma brucei.
RT Comparison of the glycosomal and cytosolic isoenzymes.";

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X83099; CAA58155.1; -
DR EMBL; Z72905; CAA97130.1; -
DR PIR; S55978; S55978.
DR SGD; S0003352; SEC35.
DR GO; GO:0017119; C:Golgi transport complex; IPI.
DR GO; GO:0000300; C:peripheral membrane protein of membrane fra. .; IDA.
DR GO; GO:0005625; C:soluble fraction; IDA.
DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0006888; P:ER to Golgi transport; IGI.
DR GO; GO:0000301; P:retrograde (vesicle recycling within Golgi) . .; IMP.
KW Transport; Protein transport; Golgi stack; Membrane.
KW SEQUENCE 275 AA; 31799 MW; AA102D086FF3PAD7 CRC64;

Query Match 38.7%; Score 46; DB 1; Length 275;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 7 ILEAMSDNTPYLH 20
|||:|||||:
DB 217 ILESCDSNSPYIH 230

RESULT 14

G3P CHLPN STANDARD; PRT; 335 AA.
ID -G3P CHLPN
AC Q9Z7T0; Q9Z0H7;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR GAPB OR CPN0624 OR CP0123
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHL029;
RC MEDLINE=9920606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RC MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uitterlind T., Berris K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwynn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mofn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RC MEDLINE=2030349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CHL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE01647; AAD18763.1; -
DR EMBL; AE002173; AAB38006.1; -
DR EMBL; AP002547; BAA98831.1; -
DR PIR; B72053; B72053.
DR PIR; B86568; B86568.
DR HSSP; P06977; IGAD.
DR PDBI-2DPAGE; Q9Z7T0; -
DR TIGR; CP0123; -
DR InterPro; IPR000173; GAP dehydrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF02800; spdh.C.1.
DR Pfam; PF02800; spdh.C.1.
DR PRINTS; PR00078; G3PDHGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I.1.
DR PROSITE; PS00071; GAPDH.1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
KW GLYCERALDEHYDE 3-PHOSPHATE (BY
KW BINDING 151 151
FT ACT_SITE 178 178
FT FT
FT ACTIVITYES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 335 AA; 36837 MW; C86D0E3AD1ADDF8E CRC64;

Query Match 38.7%; Score 46; DB 1; Length 335;
Best Local Similarity 52.9%; Pred. No. 8.2;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 2 NDHLSILEAMSDNTPY 18
|||:|||||:
DB 303 NDHFPFLVAMYNENERGY 319

RESULT 15

PCHA_PSEAE STANDARD; PRT; 476 AA.
ID -PCHA_PSEAE
AC Q51506;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salicylate biosynthesis isochorismate synthase (EC 5.4.99.6)
DE (isochorismate mutase).
GN PCHA OR PA4231.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RC MEDLINE=9608693; PubMed=7500944;
RA Serrino L., Reilmann C., Baur H., Beyeler M., Visca P., Haas D.;
RT "Structural genes for salicylate biosynthesis from chorismate in
RT Pseudomonas aeruginosa.";
RL Mol. Gen. Genet. 249:217-228(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-10
Perfect score: 119
Sequence: 1 ANDHLSIEAMSNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	90.8	1590	2	Q59983 streptococc
2	108	90.8	1590	2	Q59983 streptococc
3	108	90.8	1590	2	Q59983 streptococc
4	108	90.8	1590	2	Q59983 streptococc
5	108	90.8	1590	2	Q59983 streptococc
6	108	90.8	1590	2	Q59983 streptococc
7	108	90.8	1590	2	Q59983 streptococc
8	108	90.8	1590	2	Q59983 streptococc
9	108	90.8	1590	2	Q59983 streptococc
10	108	90.8	1590	2	Q59983 streptococc
11	108	90.8	1590	2	Q59983 streptococc
12	108	90.8	1590	2	Q59983 streptococc
13	108	90.8	1590	2	Q59983 streptococc
14	108	90.8	1590	2	Q59983 streptococc
15	108	90.8	1590	2	Q59983 streptococc
16	108	90.8	1590	2	Q59983 streptococc

17	62	52.1	1512	2	Q9WXJ5	Q9WXJ5 streptococc
18	60	50.4	2057	2	Q9RE05	Q9RE05 leucostoc
19	58	48.7	1518	2	Q00600	Q00600 streptococc
20	57	47.9	1599	2	Q00599	Q00599 streptococc
21	55.5	46.6	1592	10	Q8RWY3	Q8RWY3 allium cepa
22	52	43.7	4848	2	Q07944	Q07944 streptococc
23	51	42.9	424	10	Q94U07	Q94U07 arabidopsis
24	51	42.9	435	10	Q94U08	Q94U08 arabidopsis
25	50	42.0	305	10	Q8LDL1	Q8LDL1 arabidopsis
26	50	42.0	339	10	Q94979	Q94979 arabidopsis
27	50	42.0	347	10	Q94Y92	Q94Y92 arabidopsis
28	49.5	41.6	55	10	Q94FX3	Q94FX3 zea mays
29	49.5	41.6	55	10	Q93VW4	Q93VW4 zea mays
30	49.5	41.6	55	10	Q93WH7	Q93WH7 zea mays
31	49.5	41.6	56	10	Q94FX8	Q94FX8 zea mays
32	49.5	41.6	129	10	Q94FX6	Q94FX6 zea mays
33	49.5	41.6	547	10	Q8LST4	Q8LST4 sorghum bic
34	49.5	41.6	549	10	Q94G64	Q94G64 zea mays
35	49.5	41.6	549	10	Q43274	Q43274 zea mays
36	49.5	41.6	601	10	Q05566	Q05566 mycobacteri
37	49	41.2	797	16	Q98P32	Q98P32 rhizobium l
38	48.5	40.8	283	5	Q9VLP7	Q9VLP7 drosophila
39	48.5	40.8	283	5	Q8WSF4	Q8WSF4 drosophila
40	48	40.3	490	16	Q06579	Q06579 mycobacteri
41	47.5	39.9	214	11	Q8C1P7	Q8C1P7 mus musculu
42	47.5	39.9	519	10	Q92UB6	Q92UB6 arabidopsis
43	47.5	39.9	534	10	Q8S528	Q8S528 arabidopsis
44	47.5	39.9	534	10	Q94C67	Q94C67 arabidopsis
45	47.5	39.9	549	10	Q91LR2	Q91LR2 oryza sativ

ALIGNMENTS

RESULT 1
ID Q59983 PRELIMINARY; PRT; 1590 AA.
AC Q59983;
DC 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 02, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5).
GN Grp1.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OM2176;
RX MEDLINE=9416405; PubMed=8312602;
RA Sato S., Inoue M., Handa N., Aizawa Y., Isebe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
RT Streptococcus sobrinus.";
RL DNA Seq. 4:19-27(1993).
DR EMBL; D13858; BAA02976.1; -;
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF03324; Glyco_hydro_70; 1.
DR Pfam; PF03324; Glyco_hydro_70; 1.
FT SIGNAL 1
FT CHAIN 39 1590
SQ SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
Best Local Similarity 85.7%; Pred. No. 5.5e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ANDHLSIEAMSNDPTPLHD 21
Db 477 ANNHSIVIEAMSNDPTPLHD 497

RESULT 2

Q55263 PRELIMINARY; PRT; 1590 AA.
ID Q55263;
AC Q55263;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE GTF-1.
GN GLUCOSYLTRANSFERASE.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1310;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478.";
RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL; D63570; BAA09792.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 15.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
Best Local Similarity 85.7%; Pred. No. 5.5e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDTPYLHD 21
DB 477 ANNHSLTEAWSDNDTPYLHD 497

RESULT 3

Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
ID Q9WXJ4;
AC Q9WXJ4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE GTF-8.
GN Streptococcus criceti.
OS Streptococcus criceti.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1333;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M.; Fukui K.; Miyagi A.;
RT "Sequence analysis of the gene encoding a glucosyltransferase (gifs and gtf8) genes.";
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB026123; BAA77236.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 65.5%; Score 78; DB 2; Length 1338;
Best Local Similarity 71.4%; Pred. No. 0.0025;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDTPYLHD 21
DB 437 AIDHSLTEAWSGNDNDYVKD 457

RESULT 4

Q9LCJ7 PRELIMINARY; PRT; 1016 AA.
ID Q9LCJ7;
AC Q9LCJ7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OC NCBI_TaxID=1245;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512P;
RX MEDLINE=20169623; PubMed=10705445;
RA Funari K.; Mizuno K.; Takahara H.; Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc mesenteroides NRRL B-512P.";
RL BioSci. Biotechnol. Biochem. 64:29-38(2000).
DR EMBL; AB020020; BAA90527.1; -.
DR HSSP; P06278; 1VUS.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFD813CCCB47 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
Best Local Similarity 61.9%; Pred. No. 0.0037;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDTPYLHD 21
DB 548 SNOHSLTEAWSDNDTPYLHD 568

RESULT 5

Q9L466 PRELIMINARY; PRT; 1477 AA.
ID Q9L466;
AC Q9L466;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSKC.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OC NCBI_TaxID=1245;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A.; Renaud-Simeon M.; Pizzut S.; Sarcabal P.;
RA Willenot R.M.; Monsen P.;
RT "Sequence analysis of the gene encoding a dextranucrase, a sucrose
glycosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ250172; CAB76565.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFC8B31 CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1477;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDTPYLHD 21
DB 532 ANGHSLTEAWSDNDTPYLHD 552

```

RESULT 6
O9EZHS PRELIMINARY; PRT; 1508 AA.
ID O9EZHS
AC O9EZHS
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase Dszb742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF294469; AAG38021.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; 1.E2FCFA0B7AA4F3A CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSLEAWSNDPTPLHD 21
Db 563 ANQHLSTLEDSHNDPEYKXD 583

RESULT 7
O52224 PRELIMINARY; PRT; 1508 AA.
ID O52224
AC O52224
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRL B-1299;
RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSLEAWSNDPTPLHD 21
Db 563 ANQHLSTLEDSHNDPEYKXD 583

RESULT 8
O9LCH3 PRELIMINARY; PRT; 1575 AA.
ID O9LCH3
AC O9LCH3
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10766934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BAA95201.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2B543 CRC64;

Query Match 61.3%; Score 73; DB 2; Length 1575;
Best Local Similarity 77.8%; Pred. No. 0.018;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 HLSLEAWSNDPTPLHD 21
Db 549 HLSLEAWSNDPTPLHD 566

RESULT 9
O54178 PRELIMINARY; PRT; 1577 AA.
ID O54178
AC O54178
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgs, which regulates expression of
RT glucosyltransferase and influences the Spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC43483.1; -.
DR EMBL; M89776; AAA26969.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase.

```

SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match Best Local Similarity 61.3%; Score 73; DB 2; Length 1577;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 HSLIIEAWSNDNDTPYLHD 21
DB 551 HSLIIEAWSNDNDPPYXND 568

RESULT 10

Q9ZAR4 PRELIMINARY; PRT; 1527 AA.

AC Q9ZAR4; 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE Dextranucrase.

OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA Bhatnagar R., Singh D.K.S.;
RT "Cloning and Molecular Characterization of Dextranucrase Gene from

RT Leuconostoc mesenteroides NRRL B-512F.";
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.

DR EMBL; U81374; AAD10952.1; -
DR InterPro; IPR002479; CW binding.

DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.

SQ SEQUENCE 1527 AA; 169709 MW; IDFAFA237C743398 CRC64;

Query Match Best Local Similarity 59.7%; Score 71; DB 2; Length 1527;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSIEAWSNDNDTPYLHD 21
DB 581 ANOHLSTLEDSHNDPLVYTD 601

RESULT 11

Q8KRE1 PRELIMINARY; PRT; 1527 AA.

AC Q8KRE1; 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Dextranucrase Dsrd (EC 2.4.1.5).

OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA Neubauer H., Bauche A., Mollet B.;
RT "Isolation and characterization of the dextranucrase Dsrd of

RT Leuconostoc mesenteroides loc4.";
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; AY017384; AAC61158.1; -
DR InterPro; IPR002479; CW_binding.

DR Pfam; PF01473; CW_binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.

DR Transferase; Glycosyltransferase.
FT NON TER 1527 AA; 169835 MW; F9D0DE220BD89668 CRC64;

Query Match Best Local Similarity 59.7%; Score 71; DB 2; Length 1527;

Best Local Similarity 66.7%; Pred. No. 0.037;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSIEAWSNDNDTPYLHD 21
DB 581 ANOHLSTLEDSHNDPLVYTD 601

RESULT 12

Q8G9Q2 PRELIMINARY; PRT; 2835 AA.

AC Q8G9Q2; 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Dextranucrase (EC 2.4.1.5) (Fragment).

OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA Boronnet S., Dolis-Lafargue M., Fabre E., Pizut S., Remaud-Simeon M.,
RT "Molecular Characterisation of DSR-E, an alpha-1,2 linkage

RT synthesising dextranucrase with two catalytic domains.";
RL J. Bacteriol. 184:5753-5761 (2002).

DR EMBL; AJ430204; CAD22883.1; -
DR Transferase; Glycosyltransferase.

FT NON TER 1
SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match Best Local Similarity 58.0%; Score 69; DB 2; Length 2835;

Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 ANDHLSIEAWSNDNDTPYL 19
DB 557 ANNHISLIEWDNDNDAYT 575

RESULT 13

Q88542 PRELIMINARY; PRT; 1449 AA.

AC Q88542; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE Glucosyltransferase N (Fragment).

OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OX NCBI_TaxID=1304;

RN SEQUENCE FROM N.A.

RA Jaffe R.I.;
RT "Streptococcus salivarius VI477 gtfN.";

RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -

DR InterPro; IPR002479; CW_binding.
DR Pfam; PF01473; CW_binding_1; 8.

DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase.

FT NON TER 1449
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BBB CRC64;

Query Match Best Local Similarity 55.5%; Score 66; DB 2; Length 1449;

Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAMSDNDTPYLHD 21
 |||||
 DB 539 HLSTLEAMSHNDAYNED 556

RESULT 14
 ID Q55264 PRELIMINARY; PRT; 1449 AA.
 AC Q55264;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTFU.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RL Infect. Immun. 63:609-621(1995).
 DR EMBL: L35495; AAC41412.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 FT CHAIN 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MM; D622F07306E86A46 CRC64;

Query Match 55.5%; Score 66; DB 2; Length 1449;
 Best Local Similarity 72.2%; Pred. No. 0.21;
 Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAMSDNDTPYLHD 21
 |||||
 DB 539 HLSTLEAMSHNDAYNED 556

RESULT 15
 ID Q8KZL5 PRELIMINARY; PRT; 1554 AA.
 AC Q8KZL5;
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 sobrinus gtfu gene that produces a highly branched water-soluble
 glucan.";
 RT Blochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL: AB089438; BAC07265.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 6.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferase.

SEQ SEQUENCE 1554 AA; 171676 MM; 6981BCC1DAE24A73 CRC64;
 Query Match 52.9%; Score 63; DB 2; Length 1554;
 Best Local Similarity 66.7%; Pred. No. 0.69;
 Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAMSDNDTPYLHD 21
 |||||
 DB 488 HLSTLEAMSLNDNQYNE 505

Search completed: November 13, 2003, 09:44:01
 UOD time : 31.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 / Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-10
Perfect score: 119
Sequence: 1 ANDHSLTLEWSDNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1475	3	US-09-007-999-2
2	119	100.0	1475	3	US-09-210-361-2
3	119	100.0	1475	3	US-09-740-274-2
4	110	92.4	1375	3	US-09-210-361-4
5	110	92.4	1375	4	US-09-740-274-4
6	75	63.0	545	4	US-09-604-957-4
7	75	63.0	1430	3	US-09-008-172-2
8	75	63.0	1430	3	US-09-210-361-6
9	75	63.0	1430	4	US-09-740-274-6
10	71	59.7	523	4	US-09-604-957-5
11	63	52.9	1577	2	US-08-793-824-2
12	60	50.4	584	4	US-09-604-957-6
13	60	50.4	2057	4	US-09-499-203-2
14	49	41.2	535	4	US-09-604-957-7
15	49	41.2	1278	4	US-09-604-957-3
16	46	38.7	336	4	US-09-198-452A-662
17	46	38.7	596	4	US-09-252-991A-31255
18	45	37.8	286	4	US-09-154-750A-80
19	45	37.8	385	1	US-08-361-920-23
20	45	37.8	385	1	US-08-479-939-23
21	45	37.8	385	1	US-08-483-432-23
22	44.5	37.4	130	4	US-09-328-352-4952
23	44.5	37.4	236	4	US-09-252-991A-31380
24	44	37.0	234	4	US-09-134-001C-4514
25	44	37.0	384	3	US-09-311-170-2
26	44	37.0	647	4	US-09-134-001C-5458
27	44	37.0	749	4	US-09-562-737-96

ALIGNMENTS

28	43	36.1	2465	2	US-08-596-291-3	Sequence 3, Appli
29	43	36.1	2465	3	US-09-100-804-3	Sequence 3, Appli
30	43	36.1	2465	3	US-09-080-855-12	Sequence 12, Appli
31	43	36.1	2466	4	US-09-566-076-12	Sequence 12, Appli
32	43	36.1	2466	5	PCN-US94-09943-2	Sequence 2, Appli
33	43	36.1	2485	5	US-09-290-640-46	Sequence 46, Appli
34	42	35.3	211	4	US-09-107-532A-5029	Sequence 5029, Ap
35	42	35.3	259	1	US-08-277-231A-3	Sequence 3, Appli
36	42	35.3	259	2	US-08-473-750-6	Sequence 6, Appli
37	42	35.3	259	4	US-08-477-326-6	Sequence 23168, A
38	42	35.3	288	4	US-09-252-991A-23168	Sequence 16, Appli
39	42	35.3	989	2	US-08-070-301-16	Sequence 30731, A
40	42	35.3	1428	4	US-09-252-991A-30731	Sequence 28000, A
41	41.5	34.9	454	4	US-09-252-991A-28000	Sequence 10, Appli
42	41.5	34.9	1048	4	US-09-171-699-10	Sequence 22373, A
43	41	34.5	244	4	US-09-252-991A-22373	Sequence 38, Appli
44	41	34.5	270	4	US-09-323-872A-38	Sequence 7769, Ap
45	41	34.5	296	4	US-09-328-352-7769	

RESULT 1
US-09-007-999-2
Sequence 2, Application US/09007999
Patent No. 6087559
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified starch and
FILE REFERENCE: 0356D
CURRENT APPLICATION NUMBER: US/09/007,999
CURRENT FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTLEWSDNDPTPLHD 21
Db 481 ANDHSLTLEWSDNDPTPLHD 501
RESULT 2
US-09-210-361-2
Sequence 2, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172

```
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2
```

```
Query Match          100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ANDHLSILEAMSDNDTPYLHD 21
      |||||
Db      481 ANDHLSILEAMSDNDTPYLHD 501
```

```
RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2
```

```
Query Match          100.0%; Score 119; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ANDHLSILEAMSDNDTPYLHD 21
      |||||
Db      481 ANDHLSILEAMSDNDTPYLHD 501
```

```
RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4
```

```
Query Match          92.4%; Score 110; DB 3; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ANDHLSILEAMSDNDTPYLHD 21
      |||||
Db      507 ANDHLSILEAMSYNDTPYLHD 527
```

```
RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4
```

```
Query Match          92.4%; Score 110; DB 4; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ANDHLSILEAMSDNDTPYLHD 21
      |||||
Db      507 ANDHLSILEAMSYNDTPYLHD 527
```

```
RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
```

```
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LOBBERT
APPLICANT: RAHMOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 545
TYPE: PR1
ORGANISM: Streptococcus mutans
US-09-604-957-4
```

```
Query Match      63.0%; Score 75; DB 4; Length 545;
Best Local Similarity 71.4%; Pred. No. 0.00043;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAMSDNDTPYLHD 21
DB 75 AINHLSTLEAMSDNDPQYNKD 95
```

```
RESULT 7
US-09-008-172-2
```

```
Sequence 2, Application US/09008172
Patent No. 6127602
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: 0358D
CURRENT APPLICATION NUMBER: US/09/008,172
CURRENT FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1430
TYPE: PR1
ORGANISM: Streptococcus mutans
US-09-008-172-2
```

```
Query Match      63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAMSDNDTPYLHD 21
DB 495 AINHLSTLEAMSDNDPQYNKD 515
```

```
RESULT 8
US-09-210-361-6
```

```
Sequence 6, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
```

```
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PR1
ORGANISM: streptococcus mutans
US-09-210-361-6
```

```
Query Match      63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAMSDNDTPYLHD 21
DB 495 AINHLSTLEAMSDNDPQYNKD 515
```

```
RESULT 9
US-09-740-274-6
```

```
Sequence 6, Application US/09740274
Patent No. 6465203
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PR1
ORGANISM: streptococcus mutans
US-09-740-274-6
```

```
Query Match      63.0%; Score 75; DB 4; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ANDHSLTEAMSDNDTPYLHD 21
DB 495 AINHLSTLEAMSDNDPQYNKD 515
```

```
RESULT 10
US-09-604-957-5
```

```
Sequence 5, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LOBBERT
```

APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 523
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 59.7%; Score 71; DB 4; Length 523;
Best Local Similarity 66.7%; Pred. No. 0.0018; 6; Indels 0; Gaps 0;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAMSNDNTPYLHD 21
Db 75 ANKHSTLEDMWNGKDPQYVD 95

RESULT 11
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of Plants to
TITLE OF INVENTION: Increase Stored Carbohydrates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Griffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AT PW7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.14; 5; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSTLEAMSNDNTPYLHD 21
Db 591 HSTLEAMSNDNTPYLHD 608

RESULT 12
US-09-604-957-6
Sequence 6, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKUIZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 584
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 50.4%; Score 60; DB 4; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.13; 6; Indels 0; Gaps 0;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAMSNDNTPYLH 20
Db 75 ANKHSTLEDMWNGKDPQYVN 94

RESULT 13
US-09-499-203-2
Sequence 2, Application US/09499203
Patent No. 6570065
GENERAL INFORMATION:
APPLICANT: KOSSMANN, Jens
APPLICANT: WELSH, Thomas
APPLICANT: QUENZ, Martin
APPLICANT: KNUTH, Karola
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
FILE REFERENCE: 147-196P
CURRENT APPLICATION NUMBER: US/09/499,203
CURRENT FILING DATE: 2000-02-08
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2057
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 50.4%; Score 60; DB 4; Length 2057;
Best Local Similarity 55.0%; Pred. No. 0.57; 6; Indels 0; Gaps 0;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAMSNDNTPYLH 20
Db 665 ANKHSTLEDMWNGKDPQYVN 684

RESULT 14
US-09-604-957-7
Sequence 7, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:


```

1  APPLICANT: VAN GEEL, SCHOTTEN GERRITDINA HENDRIKA
2  APPLICANT: DICKHUIZEN, LUBBERT
3  APPLICANT: RAHAOUI, HAKIM
4  APPLICANT: LEER, ROBERT-JAN
5  TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
6  FILE REFERENCE: BO 43388
7  CURRENT APPLICATION NUMBER: US/09/604,957
8  CURRENT FILING DATE: 2000-06-28
9  PRIOR APPLICATION NUMBER: 00201871.1
10 PRIOR FILING DATE: 2000-05-25
11 NUMBER OF SEQ ID NOS: 17
12 SOFTWARE: PatentIn Ver. 2.1
13 SEQ ID NO 7
14 LENGTH: 535
15 TYPE: PR1
16 ORGANISM: Lactobacillus reuteri
17 US-09-604-957-7

```

Query Match	41.2%;	Score 49;	DB 4;	Length 535;
Best Local Similarity	40.0%;	Pred. No. 7;		
Matches	8;	Conservative	5;	Mismatches 7;
				Indels 0;
				Gaps 0;

QY 1 ANDHLSILEAWSNDNTPYLH 20
:|::||| |: | :
Db 74 SNKHINILEDWNHADPEYFN 93

```

, RESULT 15
, US-09-604-957-3
, Sequence 3, Application US/09604957
, Patent No. 6486314
, GENERAL INFORMATION:
, APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
, APPLICANT: DIJKHUIZEN, LUBBERT
, APPLICANT: RAHNOUI, HAKIM
, APPLICANT: LEER, ROBERT-JAN
, TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
, FILE REFERENCE: BO 4338
, CURRENT APPLICATION NUMBER: US/09/604,957
, CURRENT FILING DATE: 2000-06-28
, PRIOR APPLICATION NUMBER: 00201871.1
, PRIOR FILING DATE: 2000-05-25
, NUMBER OF SEQ ID NOS: 17
, SOFTWARE: PatentIn Ver. 2.1
, SEQ ID NO 3
, LENGTH: 1278
, TYPE: PR1
, ORGANISM: Lactobacillus reuteri
, US-09-604-957-3

```

Query Match	41.2%	Score 49;	DB 4;	Length 1278;
Best Local Similarity	40.0%;	Pred. No. 20;		
Matches	8;	Conservative	5;	Mismatches 7; Indels 0; Gaps 0

```
QY      1 ANDHLSILEAWSNDNTPYLH 20
      :|::|||:|:
Db      550 SNKHINILLEWNNHADPEYFN 569
```

Search completed: November 13, 2003, 09:47:59
Job time : 14.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-10

Perfect score: 119

Sequence: 1 ANDHLSITLAWSDNDPTLYLHD 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1475	9	US-09-740-274-2
2	110	92.4	1375	9	US-09-740-274-4
3	75	63.0	545	10	US-09-995-749A-10
4	75	63.0	1430	9	US-09-740-274-6
5	71	59.7	522	10	US-09-995-749A-11
6	60	50.4	584	10	US-09-995-749A-12
7	51	42.9	502	12	US-10-009-823A-1
8	49	41.2	302	9	US-09-815-242-14004
9	49	41.2	535	10	US-09-995-749A-13
10	49	41.2	1781	10	US-09-995-749A-2
11	47.5	39.9	534	10	US-09-344-882-24
12	47.5	39.9	534	15	US-10-293-865-24
13	47	39.5	15	12	US-10-295-693-90
14	47	39.5	15	15	US-10-067-649-90
15	46.5	39.1	1604	9	US-09-888-615-73

16	46	38.7	199	10	US-09-738-626-4571	Sequence 4571, Ap
17	46	38.7	476	11	US-09-908-299-9	Sequence 9, Appl
18	46	38.7	539	15	US-10-230-026-34	Sequence 34, Appl
19	45.5	38.2	156	12	US-10-001-245-125	Sequence 125, App
20	45	37.8	268	15	US-10-156-761-13823	Sequence 167, App
21	44.5	37.4	210	15	US-10-127-032-167	Sequence 9158, Ap
22	44	37.0	216	15	US-10-156-761-9158	Sequence 11634, A
23	44	37.0	332	9	US-09-815-242-11634	Sequence 10188, A
24	44	37.0	331	9	US-09-815-242-10188	Sequence 13981, A
25	44	37.0	334	9	US-09-815-242-13981	Sequence 96, Appl
26	44	37.0	749	15	US-10-211-962-96	Sequence 7, Appl
27	44	37.0	816	15	US-10-080-114A-7	Sequence 126, App
28	43.5	36.6	155	12	US-10-001-245-126	Sequence 3308, Ap
29	43	36.1	41	11	US-09-764-891-3308	Sequence 18, Appl
30	43	36.1	311	11	US-09-878-781-18	Sequence 9, Appl
31	43	36.1	358	11	US-09-095-478-9	Sequence 7, Appl
32	43	36.1	381	11	US-09-095-478-7	Sequence 7562, Ap
33	43	36.1	715	15	US-10-156-761-7562	Sequence 56, Appl
34	43	36.1	1267	15	US-10-059-585-56	Sequence 12, Appl
35	43	36.1	2466	12	US-10-177-980-12	Sequence 46, Appl
36	43	36.1	2485	9	US-09-802-669-46	Sequence 2, Appl
37	43	36.1	2654	15	US-10-227-610-2	Sequence 10190, A
38	42.5	35.7	491	9	US-09-815-242-10190	Sequence 13756, A
39	42.5	35.7	491	9	US-09-815-242-13756	Sequence 40, Appl
40	42	35.3	31	15	US-10-092-908-40	Sequence 14, Appl
41	42	35.3	41	15	US-10-092-908-41	Sequence 26, Appl
42	42	35.3	95	9	US-09-945-301-14	Sequence 8855, Ap
43	42	35.3	298	11	US-09-988-0678-26	Sequence 8339, Ap
44	42	35.3	348	10	US-09-738-626-6855	
45	42	35.3	484	15	US-10-156-761-8339	

ALIGNMENTS

RESULT 1

US-09-740-274-2

Sequence 2, Application US/09740274

Patent No. US20020031826A1

GENERAL INFORMATION:

APPLICANT: Nicholse, Scott E.

TITLE OF INVENTION: Glucan-containing Compositions and Paper

FILE REFERENCE: 0357CRD

CURRENT APPLICATION NUMBER: US/09/740, 274

CURRENT FILING DATE: 2000-12-19

PRIOR APPLICATION NUMBER: 09/210,361

PRIOR FILING DATE: 1998-12-11

PRIOR APPLICATION NUMBER: 09/007, 999

PRIOR FILING DATE: 1998-01-16

PRIOR APPLICATION NUMBER: 08/478, 704

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: 09/006, 172

PRIOR FILING DATE: 1998-01-16

PRIOR APPLICATION NUMBER: 08/482, 711

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 1475

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-740-274-2

Query Match 100.0% ; Score 119, DB 9; Length 1475;
Best Local Similarity 100.0% ; Pred. No. 1.8e-09;
Matches 21, Conservative 0; Mismatches 0; Gaps 0;

QY 1 ANDHLSITLAWSDNDPTLYLHD 21
|||||

Db 481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 92.4%; Score 110; DB 9; Length 1375;

Best Local Similarity 95.2%; Pred. No. 4e-08; 1; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 1;

Qy 1 ANDHLSILEAMSDNDTPYLHD 21

Db 507 ANDHLSILEAMSDNDTPYLHD 527

RESULT 3

US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 63.0%; Score 75; DB 10; Length 545;

Best Local Similarity 71.4%; Pred. No. 0.0033; 5; Indels 0; Gaps 0;
Matches 15; Conservative 1; Mismatches 5;

Qy 1 ANDHLSILEAMSDNDTPYLHD 21

Db 75 AINHLSILEAMSDNDPQYNKD 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 63.0%; Score 75; DB 9; Length 1430;

Best Local Similarity 71.4%; Pred. No. 0.0097; 5; Indels 0; Gaps 0;
Matches 15; Conservative 1; Mismatches 5;

Qy 1 ANDHLSILEAMSDNDTPYLHD 21

Db 495 AINHLSILEAMSDNDPQYNKD 515

RESULT 5

US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 59.7%; Score 71; DB 10; Length 522;

Best Local Similarity 66.7%; Pred. No. 0.013;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 21
Db 75 ANQHLSILEDWMSHNDPLVYTD 95

RESULT 6
US-09-995-749A-12

; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 50.4%; Score 60; DB 10; Length 584;

Best Local Similarity 55.0%; Pred. No. 0.72;

Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 20
Db 75 ANQHLSILEDWMSHNDPLVYTD 94

RESULT 7
US-10-009-823A-1

; Sequence 1, Application US/10009823A
; Publication No. US20030157120A1
; GENERAL INFORMATION:
; APPLICANT: Fanuccio, Michael
; APPLICANT: Rosey, Everett Lee
; APPLICANT: Sinistraj, Meri
; APPLICANT: Hasse, Detlef
; APPLICANT: Parsons, Jim
; APPLICANT: Ankenbauer, Robert G.
; TITLE OF INVENTION: LAWSONIA DERIVED GENE AND RELATED FLG
; FILE REFERENCE: POLYPEPTIDES, PEPTIDES AND PROTEINS AND THEIR USES
; CURRENT APPLICATION NUMBER: US/10/009,823A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00437
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/133,973
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-10-009-823A-1

Query Match 42.9%; Score 51; DB 12; Length 502;

Best Local Similarity 47.1%; Pred. No. 14;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTP 17
Db 185 ANPYFALIESWKGNGTP 201

RESULT 8
US-09-815-242-14004

; Sequence 14004, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14004
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14004

Query Match 41.2%; Score 49; DB 9; Length 302;

Best Local Similarity 50.0%; Pred. No. 17;

Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDTPYLHD 21
Db 267 NDRLVIOEVLEASDIPYRHD 286

RESULT 9
US-09-995-749A-13

; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 41.2%; Score 49; DB 10; Length 535;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHSLLEAMSDNDTPYLH 20
DB 74 SNKHINILEDWNADEPEYFN 93

RESULT 10
US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHACUI, HAKIM
; APPLICANT: LEBER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 41.2%; Score 49; DB 10; Length 1781;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHSLLEAMSDNDTPYLH 20
DB 1053 SNKHINILEDWNADEPEYFN 1072

RESULT 11
US-09-344-882-24
; Sequence 24, Application US/09344882
; Patent No. US20020162137A1
; GENERAL INFORMATION:
; APPLICANT: Nikolaou, Basil J
; APPLICANT: Wurttele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allied, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 24
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-24

Query Match 39.9%; Score 47.5; DB 10; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLLEAMSDNDTPY 18
DB 132 NDEIAALETW-DNGKPY 147

RESULT 12
US-10-293-865-24
; Sequence 24, Application US/10293865
; Publication No. US20030106090A1
; GENERAL INFORMATION:
; APPLICANT: Nikolaou, Basil J
; APPLICANT: Wurttele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allied, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; FILE REFERENCE: 217113
; CURRENT APPLICATION NUMBER: US/10/293,865
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 24
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-10-293-865-24

Query Match 39.9%; Score 47.5; DB 15; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLLEAMSDNDTPY 18
DB 132 NDEIAALETW-DNGKPY 147

RESULT 13
US-10-295-693-90
; Sequence 90, Application US/10295693
; Publication No. US20030198976A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBRMT14, RELATED TO THE
; FILE REFERENCE: D0118 CIP
; CURRENT APPLICATION NUMBER: US/10/295,693
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US 60/266,525
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 10/067,649
; PRIOR FILING DATE: 2002-02-05

```
;; PRIOR APPLICATION NUMBER: US 60/329,897
;; PRIOR FILING DATE: 2001-10-16
;; NUMBER OF SEQ ID NOS: 95
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 90
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Synthesized Random Peptide.
US-10-295-693-90

Query Match      39.5%; Score 47; DB 12; Length 15;
Best Local Similarity 61.5%; Pred. No. 1.2;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      8 LEAMSDNDTPYLH 20
Db      3 LEAMDLSDTPHLY 15

RESULT 14
US-10-067-649-90
; Sequence 90, Application US/10067649
; Publication No. US2003010057A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBRM14, RELATED TO T
; TITLE OF INVENTION: ORPHAN GPCR, GPR73
; FILE REFERENCE: D0118 NP
; CURRENT APPLICATION NUMBER: US/10/067,649
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,525
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/329,897
; PRIOR FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized Random Peptide.
US-10-067-649-90

Query Match      39.5%; Score 47; DB 15; Length 15;
Best Local Similarity 61.5%; Pred. No. 1.2;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      8 LEAMSDNDTPYLH 20
Db      3 LEAMDLSDTPHLY 15

RESULT 15
US-09-888-615-73
; Sequence 73, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
```

```
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 73
;; LENGTH: 1604
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-888-615-73

Query Match      39.1%; Score 46.5; DB 9; Length 1604;
Best Local Similarity 47.4%; Pred. No. 2.6e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY      5 LSILEAMSDN---DTPYLH 20
Db      294 VALLEWVKDNRRTDIPELH 312

Search completed: November 13, 2003, 10:29:01
Job time : 27.4692 secs
```


PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR N-PSDB; ABK52940.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Disclosure; Page 38-42; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFD.

XX Sequence 1430 AA;

XX Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AINHLSTLEAWSDNDPOYVND 21
 ||||||||||||||||||||
 DB 495 AINHLSTLEAWSDNDPOYVND 515

RESULT 2
 AAU98041 standard; Protein; 1430 AA.
 ID AAU98041;
 AC AAU98041;
 XX
 XX
 DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFD mutant T589D.
 DE
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.

OS Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 599 /note= "wild-type Thr substituted by Asp"
 FT
 FT
 FT
 PN US2002031826-A1.
 PN
 XX 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007899.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT
 XX
 XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sized and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

SQ Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AINHLSTLEAMSDNDPOYKND 21
 |||||
 495 AINHLSTLEAMSDNDPOYKND 515

RESULT 3
 ID AAU98042 standard; Protein; 1430 AA.
 XX AAU98042;
 AC AAU98042;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFD mutant T589E.
 DE
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 589 /note= "Wild-type Thr substituted by Glu"
 FT
 FT
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 PD
 PD 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 PA Nichols SE;
 PI
 PI WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from T448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

SQ Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AINHLSTLEAMSDNDPOYKND 21
 |||||
 495 AINHLSTLEAMSDNDPOYKND 515

RESULT 4
 ID AAU98043 standard; Protein; 1430 AA.
 XX AAU98043;
 AC AAU98043;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFD mutant N471D.
 DE
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 471 /note= "Wild-type Asn substituted by Asp"
 FT
 FT
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 PD
 PD 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 PA Nichols SE;
 PI
 PI WPI; 2002-414332/44.

CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

CC Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSDNDPOYNKD 21
 |||||
 DB 495 AINHLSTLEAWSDNDPOYNKD 515

RESULT 6
 AAU98045
 ID AAU98045 standard; Protein; 1430 AA.

XX AAU98045;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFD mutant N471D/T589E.

XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

XX amyloplast; vacuole; paper manufacture; mutant; mutuin.

OS Streptococcus mutans.

XX Synthetic.

FT Key Location/Qualifiers

FT Misc-difference 471 /note= "Wild-type Asn substituted by Asp"

FT Misc-difference 589 /note= "Wild-type Thr substituted by Glu"

XX US2002031826-A1.

PD 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0462711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/)/ NICHOLS S E.

XX Nichols SE;

XX WPI, 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful

PT as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from 148V, D457N,

CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the

CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

CC coating composition comprising a glucan produced in a plant transformed

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC the glucan is produced in the amyloplast and/or vacuole or a maize line

CC deficient in starch biosynthesis, transformed with a gene encoding a

CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful

CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing

CC conditions to produce a regenerated plant and inducing expression of the

CC polynucleotide for a time sufficient to produce the glucan in the

CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC is produced in the amyloplast of potato or the vacuole of sugar beet.

CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper

CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFD mutant of the invention.

CC Note: The present sequence is not shown in the specification but

CC was created by the indexer using the GTFD sequence appearing as AAU98029

CC and the information in claim 36.

XX Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSDNDPOYNKD 21
 |||||
 DB 495 AINHLSTLEAWSDNDPOYNKD 515

RESULT 7
 AAR91047
 ID AAR91047 standard; Protein; 1577 AA.

XX AAR91047;

DT 22-MAY-1996 (first entry)

XX Alpha-D-glucosyltransferase.

XX Alpha-D-glucosyltransferase; primer-independent; soluble glucan;

KW sucrose; transgenic plant; cloning; Escherichia coli;

KW phage lambda-C13; vector; plasmid pSG501; plasmid pSG502;

KW gene transfer; crop improvement; storage carbohydrate; pasture;

KW feedstuff; senescence; dextran; binder; food; pharmaceutical.

XX Streptococcus salivarius strain ATCC 25975.

XX W09606173-A1.

XX 29-FEB-1996.

XX 24-AUG-1995; 95WO-AU00527.

XX 24-AUG-1994; 94AU-0007643.

XX (GTFE/)/ GIFFARD P M.

XX (UACQ/)/ JACQUES N A.

XX (STMP/)/ SIMPSON C L.

XX Giffard PM, Jacques NA, Simpson CL;

XX WP1, 1996-1513/6/15.
DR N-PSDB; AAT13139.
XX
XX Plants contg. new bacterial DNA encoding glucosyl transferase
PT activity - retain higher levels of stored carbohydrate(s) in a form
PT readily digestible by ruminants
XX
XX Claim 4; Page 16-20; 31pp; English.
XX
XX The sequence represents an alpha-D-glucosyltransferase from
CC Streptococcus salivarius. The enzyme is primer-independent, and
CC produces soluble glucan from sucrose. A gene encoding the enzyme
CC may be cloned and expressed in Escherichia coli using a subclone
CC of phage lambda-C13, e.g. plasmid pSGS501 or plasmid pSGS502. The
CC DNA may also be expressed in a transgenic plant, to improve the
CC level of stored carbohydrate in a pasture plant which normally
CC contains low levels, or to prevent degradation of stored carbohydrate
CC during plant senescence. Dextran may be isolated from the plant, for
CC use as a food binder or pharmaceutical additive. Primer independence
CC ensures that the enzyme will be functional in plants. The glucan is
CC poorly degraded in plants but easily degraded by bacteria in the rumen
CC of grazing livestock.

Key	Location/Qualifiers
Misc-difference	457
US2002031826-A1.	/note= "wild-type Asp substituted by Asn"
14-MAR-2002.	
19-DEC-2000;	2000US-0740274.
11-DEC-1998;	98US-0210361.
07-JUN-1995;	95US-0218704.
07-JUN-1995;	95US-0482711.
07-JUN-1995;	95US-0485243.
16-JAN-1998;	98US-0007999.
16-JAN-1998;	98US-0008172.
20-JAN-1998;	98US-0009620.
(NICH/) NICHOLS S E.	
Nichols SE;	
WPI;	2002-414332/44.
Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions	
Claim 36; Page -; 44p; English.	
The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, a polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the expression cassette, host cell promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant, a paper sizing and/or vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a trans sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTF mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AA096027 and the information in claim 36.	

XX	Sequence	1475 AA;
XX	Query Match	64.7%; Score 75; DB 23; Length 1475;
XX	Best Local Similarity	71.4%; Pred. No. 0.0092;
XX	Matches 15; Conservative	1; Mismatches 5; Indels 0; Gaps 0;
OY	1 AINHLSTLEAWSDDNDPOYNKD 21	
DB	481 ANDHLSLEAWSDDNDPTYLHD 501	
RESULT 12		
ID	AAU98032 standard; Protein; 1475 AA.	
XX	AAU98032	
AC	AAU98032;	
XX		
DT	27-AUG-2002 (first entry)	
XX		
DE	S. mutans glucosyltransferase GTFB mutant D567N.	
XX		
KW	Glucosyltransferase; GTFB; transgenic plant; paper sizing;	
KM	coating composition; glucan; starch; latex; thermoplastic molecule;	
KW	amyloplast; vacuole; paper manufacture; mutant; muten.	
XX		
OS	Streptococcus mutans.	
XX	Synthetic.	
FN	Key	Location/Qualifiers
FT	Misc-difference 567	
FT	/note= "Wild-type Asp substituted by Thr"	
XX		
PN	US2002031826-A1.	
XX		
PD	14-MAR-2002.	
XX		
PF	19-DEC-2000; 2000US-0740274.	
XX		
XX	11-DEC-1998; 98US-0210361.	
PR	07-JUN-1995; 95US-0478704.	
PR	07-JUN-1995; 95US-0482711.	
PR	07-JUN-1995; 95US-0485243.	
PR	16-JAN-1998; 98US-0007999.	
PR	16-JAN-1998; 98US-0008172.	
PR	20-JAN-1998; 98US-0009620.	
XX		
PA	(NICH/) NICHOLS S E.	
XX		
PI	Nichols SE;	
XX		
XX	WPI; 2002-414332/44.	
DR		
XX		
PT	Glucosyltransferase B or D protein useful for producing a glucan useful	
PT	as substitutes for and additions to modified starch and latexes in	
PT	paper manufacture, comprises mutations in specific positions -	
XX		
PS	Claim 36; Page -; 44p; English.	
XX		
CC	The invention an isolated protein comprising a glucosyltransferase	
CC	(GTF) B polypeptide having changes at position from I448V, D457N,	
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,	
CC	D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779G/K1014T,	
CC	Y169A/Y170A/Y171A, and K779G or a GTF D polypeptide having	
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and	
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,	
CC	an isolated polynucleotide which encodes P1 or P2, or its complementary	
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,	
CC	an expression cassette comprising the polynucleotide operably linked to a	
CC	promoter, a vector comprising the expression cassette, host cell	
CC	introduced with the vector, a transgenic plant comprising the	
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or	
CC	coating composition comprising a glucan produced in a plant transformed	

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 64.7%; Score 75; DB 23; Length 1475;

Best Local Similarity 71.4%; Pred. No. 0.0092; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5;

QY 1 AINHSTLEAWSNDPQYKND 21

Db 481 ANDHSTLEAWSNDPQYKND 501

RESULT 13
 AAU98033 ID AAU98033 standard; Protein; 1475 AA.

XX AC AAU98033;

XX DT 27-AUG-2002 (first entry)

XX DE S. mutans glucosyltransferase GTFB mutant K1014T.

XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX KM amyloplast; vacuole; paper manufacture; mutant; muten.

XX OS Streptococcus mutans.
 XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1014 /note="Wild-type lys substituted by Thr"

XX PN US2002031826-A1.

XX PD 14-MAR-2002.

XX PF 19-DEC-2000; 2000US-0740274.

XX PR 11-DEC-1998; 98US-0210361.

XX PR 07-JUN-1995; 95US-0478704.

XX PR 07-JUN-1995; 95US-0482711.

XX PR 07-JUN-1995; 95US-0485243.

XX PR 16-JAN-1998; 98US-0007999.

XX PR 16-JAN-1998; 98US-0008172.

XX PR 20-JAN-1998; 98US-0009620.

XX PA (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
 XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX an isolated polynucleotide which encodes p1 or p2, or its complementary
 XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX an expression cassette comprising the expression cassette, host cell
 XX promoter, a vector comprising the transgenic plant comprising the
 XX introduced with the vector, a transgenic plant, a paper sizing and/or
 XX vector, a seed or tuber from the transgenic plant, a plant transformed
 XX coating composition comprising a glucan produced in a plant transformed
 XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
 XX thermoplastic molecule or their combinations and/or vacuole or a maize line
 XX deficient in starch biosynthesis, transformed with a gene encoding a
 XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 XX comprising the glucan (paper sizing/coating agent). The vector is useful
 XX for producing a glucan in a plant. The method comprises transforming a
 XX plant cell with the vector, growing the plant cell under plant growing
 XX conditions to produce a regenerated plant and inducing expression of the
 XX polynucleotide for a time sufficient to produce the glucan in the
 XX regenerated plant, where the vector contains a transit sequence from
 XX ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 XX chlorophyll AB binding protein to produce a transgenic plant, and glucan
 XX is produced in the amyloplast of potato or the vacuole of sugar beet.
 XX Glucans are useful as substitutes for and additions to modified starch
 XX and latexes in paper manufacture. Unlike prior art techniques, which
 XX require input materials that produce chemical effluents, paper
 XX manufacture utilizing the glucan produced by GTF, which utilizes
 XX biologically produced input materials, is more cost-effective and
 XX environmentally friendly. Moreover, glucans also exhibit thermoplastic
 XX properties and impart gloss to the paper during coating step.
 XX The present sequence represents a GTFB mutant of the invention.
 XX Note: The present sequence is not shown in the specification but
 XX was created by the indexer using the GTFB sequence appearing as AAU98027
 XX and the information in claim 36.

XX Sequence 1475 AA;

Query Match 64.7%; Score 75; DB 23; Length 1475;

Best Local Similarity 71.4%; Pred. No. 0.0092; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5;

QY 1 AINHSTLEAWSNDPQYKND 21

Db 481 ANDHSTLEAWSNDPQYKND 501

RESULT 14

AAU98034 ID AAU98034 standard; Protein; 1475 AA.

XX AC AAU98034;

XX DT 27-AUG-2002 (first entry)

XX DE S. mutans glucosyltransferase GTFB mutant D457N/D567T.

XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567
 PT Misc-difference /note= "Wild-type Asp substituted by Thr"
 XX US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 FT Claim 36; Page -; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX
 SQ Sequence 1475 AA;
 QY 1 AINHLSTLEAMSNDPEQYKND 21
 Db 481 ANDHLSTLEAMSNDPEYLLHD 501
 Query Match 64.7%; Score 75; DB 23; Length 1475;
 Best Local Similarity 71.4%; Pred. No. 0.0092;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 RESULT 15
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.
 XX
 AC AAU98035;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571
 FT Misc-difference /note= "Wild-type Asp substituted by Lys"
 XX
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 FT Claim 36; Page -; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary

polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing the glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTPB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTPB sequence appearing as AAU98027 and the information in claim 36.

XX
SQ Sequence 1475 AA;

Query Match 64.7%; Score 75; DB 23; Length 1475;
Beet Local Similarity 71.4%; Pred. No. 0.0092;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
| : ||||| ||||| |||||
Db 481 ANDHSLTEAWSNDNDPTPLHD 501

Search completed: November 13, 2003, 09:38:26
Job time : 38.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-11
Perfect score: 116
Sequence: 1 AINHSLILEAMSDNDPQYKND 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	1431	2 A45866	dextranucrase (EC
2	90	77.6	1577	2 T30858	glucosyltransferas
3	85	73.3	1449	2 T30857	glucosyltransferas
4	85	73.3	1449	2 T30852	glucosyltransferas
5	81	69.8	1365	2 A41483	glucosyltransferas
6	78	67.2	1508	2 T31098	probable dextranu
7	75	64.7	1475	2 B31135	gtfB protein precu
8	74	63.8	1592	2 A38175	glucosyltransferas
9	68	58.6	1518	2 A44811	glucosyltransferas
10	67	57.8	1599	2 S22737	glucosyltransferas
11	66	56.9	1375	2 JT0345	dextranucrase (EC
12	51	44.0	429	2 AH0630	4-hydroxyphenylace
13	49	42.2	632	2 T46504	hypothetical prote
14	47.5	40.9	369	2 AE2493	hypothetical prote
15	47	40.5	175	2 C86205	hypothetical prote
16	47	40.5	403	2 B70961	probable esterase
17	47	40.5	418	1 FOXRL2	sigma 2 protein -
18	47	40.5	445	2 H75360	cyclochrome P450 -
19	47	40.5	491	2 A86824	sensor protein kin
20	46	39.7	418	1 FOXRL5	sigma 2 protein -
21	46	39.7	420	2 T05877	hypothetical prote
22	46	39.7	623	2 T35377	probable membrane
23	46	39.7	703	2 D70178	PTS system, fructo
24	46	39.7	703	2 A48764	calpain (EC 3.4.22
25	45.5	39.2	250	2 G72495	probable polysulfi
26	45	38.8	194	2 H72037	conserved hypothet
27	45	38.8	194	2 C86586	CT647 hypothetical
28	45	38.8	249	2 F95957	probable transcrip
29	45	38.8	344	2 T21604	hypothetical prote

30	45	38.8	440	2 C84265	adenylosuccinate s
31	45	38.8	504	2 S51942	prunin 2 precursor
32	45	38.8	637	2 T00548	hypothetical prote
33	45	38.8	644	2 A97268	methionyl-tRNA syn
34	45	38.8	657	2 S25184	cspl protein - Cor
35	45	38.8	1475	2 T29809	hypothetical prote
36	45	38.8	2358	2 T39569	probable alpha-glu
37	45	38.8	2371	2 T43432	alpha-glucan synth
38	44.5	38.4	188	2 A75382	hypothetical prote
39	44.5	38.4	236	2 S54428	fimbrial protein h
40	44.5	38.4	241	1 S24978	conserved hypothet
41	44	37.9	122	2 AD1009	conserved hypothet
42	44	37.9	291	2 AB3235	nitrotriacetate
43	44	37.9	312	2 S67667	hypothetical prote
44	44	37.9	335	2 C85642	hypothetical prote
45	44	37.9	335	2 F90781	hypothetical prote

ALIGNMENTS

RESULT 1

A45866 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans

C/Species: Streptococcus mutans

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: A45866

R/Honda, O.; Kato, C.; Kuramitsu, H.K.

J. Gen. Microbiol. 136, 2099-2105, 1990

A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl

A/Reference number: A45866; MUID:91100958; PMID:2148600

A/Accession: A45866

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1431 <HON>

A/Cross-references: GB:M29296

C/Superfamily: cpl repeat homology

C/Keywords: glucosyltransferase; hexosyltransferase

F/181-201/Domain: cpl repeat homology <CP1>

F/1127-1146/Domain: cpl repeat homology <CP2>

F/1192-1211/Domain: cpl repeat homology <CP3>

F/1257-1276/Domain: cpl repeat homology <CP4>

F/1277-1297/Domain: cpl repeat homology <CP5>

F/1321-1340/Domain: cpl repeat homology <CP6>

F/1341-1361/Domain: cpl repeat homology <CP7>

F/1385-1404/Domain: cpl repeat homology <CP7>

Query Match 100.0%; Score 116; DB 2; Length 1431;

Best Local Similarity 100.0%; Pred. No. 2.2e-03; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0;

QY 1 AINHSLILEAMSDNDPQYKND 21

DB 495 AINHSLILEAMSDNDPQYKND 515

RESULT 2

T30858 glucosyltransferase - Streptococcus salivarius

C/Species: Streptococcus salivarius

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C/Accession: T30858

R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for phage

A/Reference number: Z20909; MUID:95122197; PMID:7822030

A/Accession: T30858

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1577 <SIM>

A/Cross-references: EMBL:L35928; NID:G662380; PID:G662381; PIDN:AA041413.1

C/Genetics:

A/Gene: gtfm

Query Match 77.6%; Score 90; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
|||
DB 588 AINHLSTLEAWSNDNDPQYKND 608

RESULT 3

T30857
glucosyltransferase - Streptococcus salivarius
C/Species: Streptococcus salivarius
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Accession: T30857
R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pr
A/Reference number: Z20909; MUID:95122197; PMID:7822030
A/Accession: T30857
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1449 <STM>
A/Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AA041412.1
C/Genetics:
A/Gene: gtfI

Query Match 73.3%; Score 85; DB 2; Length 1449;
Best Local Similarity 76.2%; Pred. No. 0.00013;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
|||
DB 536 AINHLSTLEAWSNDNDPQYKND 556

RESULT 4

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C/Species: Streptococcus salivarius
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Accession: T30552
R/Jaffe, R.I.
submitted to the EMBL Data Library, February 1998
A/Description: Streptococcus salivarius V1477 gtfN.
A/Reference number: Z20854
A/Accession: T30552
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1449 <JAF>
A/Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AA005156.1
C/Genetics:
A/Gene: gtfN

Query Match 73.3%; Score 85; DB 2; Length 1449;
Best Local Similarity 76.2%; Pred. No. 0.00013;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
|||
DB 536 AINHLSTLEAWSNDNDPQYKND 556

RESULT 5

AA1483
glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus
C/Species: Streptococcus sobrinus
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C/Accession: A41483
R/Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
Infect. Immun. 58, 2452-2458, 1990
A/Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltrans

A/Reference number: A41483; MUID:9031665; PMID:2142479
A/Accession: A41483
A/Molecule type: DNA
A/Residues: 1-1365 <GIL>
A/Cross-references: GB:M30943; NID:g153652; PIDN:AAA26898.1; PID:g153653
C/Genetics:
A/Gene: gtfS
C/Superfamily: cpl repeat homology
C/Keywords: glycosyltransferase, hexosyltransferase

Query Match 69.8%; Score 81; DB 2; Length 1365;
Best Local Similarity 76.2%; Pred. No. 0.0005;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
|||
DB 467 AINHLSTLEAWSNDNDPQYKND 487

RESULT 6

T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
C/Species: Leuconostoc mesenteroides
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C/Accession: T31098
R/Monchois, V.; Renaud-Simeon, M.; Monsau, P.; Willmet, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A/Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A/Reference number: Z20981; MUID:98164374; PMID:9503626
A/Accession: T31098
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1508 <MON>
A/Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AAB95453.1
A/Experimental source: strain NRRL B-1299
C/Genetics:
A/Gene: dsrB
C/Function:
A/Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C/Keywords: glycosyltransferase, hexosyltransferase

Query Match 67.2%; Score 78; DB 2; Length 1508;
Best Local Similarity 71.4%; Pred. No. 0.0016;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
|||
DB 563 AINHLSTLEAWSNDNDPQYKND 583

RESULT 7

B33135
gtfB protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C/Accession: B33135; A33128
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Accession: B33135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1475 <SHI>
A/Cross-references: GB:M17361; NID:g153639; PIDN:AAA88588.1; PID:g153640
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990
A/Reference number: A33128
A/Accession: A33128
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-171,173-641, 'N', 643-1475 <SH2>
A/Experimental source: strain GS-5

C:Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 64.7%; Score 75; DB 2; Length 1475;
Best Local Similarity 71.4%; Pred. No. 0.0045;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 21
DB 481 AINHLSTLEAWSNDPQYKD 501

RESULT 8
A:8175
glucosyltransferase precursor - Streptococcus sobrinus

C:Species: Streptococcus sobrinus
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1995
C:Accession: A38175
R:Abdo, H.; Msumura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
J. Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A:Reference number: A38175; MUID:91123227; PMID:1704006
A:Accession: A38175
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1592 <ABO>
A:Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g217033
C:Superfamily: cpl repeat homology
F:1093-1112/Domain: cpl repeat homology <CP1>
F:1222-1241/Domain: cpl repeat homology <CP2>
F:1287-1306/Domain: cpl repeat homology <CP3>
F:1330-1351/Domain: cpl repeat homology <CP4>
F:1352-1371/Domain: cpl repeat homology <CP5>
F:1402-1420/Domain: cpl repeat homology <CP6>
F:1465-1484/Domain: cpl repeat homology <CP7>
F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 63.8%; Score 74; DB 2; Length 1592;
Best Local Similarity 66.7%; Pred. No. 0.007;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 21
DB 477 AINHLSTLEAWSNDPQYKD 497

RESULT 9
A:44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius

C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811
A:Molecule type: DNA
A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAA77900.1; PID:g47527
A:Note: sequence extracted from NCBI backbone (NCBI:81050, NCBI:81052)
C:Genetics:
A:Gene: gtfJ
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 58.6%; Score 68; DB 2; Length 1518;
Best Local Similarity 63.2%; Pred. No. 0.056;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 19
DB 501 AINHLSTLEAWSNDPQYKD 519

RESULT 10

S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C:Accession: S22737; S28810; E44811; S22727
R:Jacques, N.
Submitted to the EMBL Data Library, March 1992
A:Reference number: S22737
A:Accession: S22737
A:Molecule type: DNA
A:Residues: 1-1599 <JAC>
A:Cross-references: EMBL:Z11873; NID:g47530; PIDN:CAA77898.1; PID:g47531
A:Experimental source: ATCC 25975
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: S28810
A:Molecule type: DNA
A:Residues: 1-51 <GIF>
A:Cross-references: EMBL:Z11873
C:Genetics:
A:Gene: gtfK
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1456-1475/Domain: cpl repeat homology <CP8>

Query Match 57.8%; Score 67; DB 2; Length 1592;
Best Local Similarity 66.7%; Pred. No. 0.084;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 18
DB 491 AINHLSTLEAWSNDPQYKD 508

RESULT 11
J070345
dextransucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)

N:Alternate names: sucrose 6-glucosyltransferase
C:Species: Streptococcus mutans
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C:Accession: J070345; C33135
R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
A:Reference number: J070345; MUID:89137980; PMID:2976010
A:Accession: J070345
A:Molecule type: DNA
A:Residues: 1-1375 <UED>
A:Experimental source: GS-5
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A:Reference number: A33135; MUID:87508013; PMID:3040685
A:Accession: C33135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <SHI>
A:Cross-references: GB:M17361
C:Genetics:
A:Gene: gtfC
C:Function:
A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C:Superfamily: cpl repeat homology

C:Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <STO>
 F:35-1375/Product: glucosyltransferase #status predicted <MNT>
 F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1253-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 56.9%; Score 66; DB 2; Length 1375;
 Best Local Similarity 66.7%; Pred. No. 0.1;
 Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPOYKND 21
 :|||:|||||:
 Db 507 ANDHLSTLEAMSYNDPIYHLD 527

RESULT 12

AH0630
 4-Hydroxyphenylacetate degradation bifunctional isomerase/decarboxylase (EC 5.3.3.-) [im
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A>Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AH0630
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 Th. T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AH0630
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-429 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08227.1; PID:g16502275; GSPDB:GN00176
 C:Genetics:
 A:Gene: hpaG
 C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 44.0%; Score 51; DB 2; Length 429;
 Best Local Similarity 47.6%; Pred. No. 5.5;
 Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 AINHLSTLEAMSD--NDPOYN 19
 :|||:|||||:
 Db 9 AINHRQLDMQAFSGPPYN 29

RESULT 13

T46504
 hypothetical protein DKFZp434M179.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46504
 R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23029
 A:Accession: T46504
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-632 <AAA>
 A:Cross-references: EMBL:AL137338
 A:Experimental source: adult testis; clone DKFZp434M179
 C:Genetics:
 A>Note: DKFZp434M179.1

Query Match 42.2%; Score 49; DB 2; Length 632;
 Best Local Similarity 52.9%; Pred. No. 17;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSTLEAMSDNDPOYKND 21
 :|||:|||||:
 Db 119 IMVLGASEPDPYKND 135

RESULT 14

hypothetical protein alr7125 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120al
 A:Accession: A2493
 C:Species: Nostoc sp. PCC 7120
 A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: A2493
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: A2493
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-369 <KOR>
 A:Cross-references: GB:BA000020; PIDN:BAF78209.1; PID:g17135663; GSPDB:GN00180
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr7125
 A:Genome: plasmid

Query Match 40.9%; Score 47.5; DB 2; Length 369;
 Best Local Similarity 40.0%; Pred. No. 16;
 Matches 10; Conservative 3; Mismatches 5; Indels 7; Gaps 1;

QY 4 HSLT-----EAMSDNDPOYKND 21
 :|||:|||||:
 Db 6 HSLTQGVVWNAWRENDIPND 30

RESULT 15

hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C86205
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Anen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talion,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; MUID:21016719; PMID:11130712
 A:Accession: C86205
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-175 <STO>
 A:Cross-references: GB:AE005172; NID:g8954041; PIDN:AAF82215.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 40.5%; Score 47; DB 2; Length 175;
 Best Local Similarity 46.2%; Pred. No. 8.2;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 INHSLTLEAMSDN 14
 :|||:|||||:
 Db 108 INHQEVYDAMSDH 120

Search completed: November 13, 2003, 09:50:25
 Job time: 14.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds

(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-11

Perfect score: 116

Sequence: 1 AINHLSTLEAMSDNDPOYNKD 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	1462	1	GTFD_STRMU P49331 streptococc
2	81	69.8	1365	1	GTFB_STRDO P29336 streptococc
3	75	64.7	1476	1	GTFB_STRMU P08987 streptococc
4	74	63.8	1592	1	GTF2_STRDO P27470 streptococc
5	74	63.8	1597	1	GTF1_STRDO P11001 streptococc
6	66	56.9	1455	1	GTRC_STRMU P13470 streptococc
7	51	44.0	429	1	HPAG_SALDU Q9rpns s 4-hydroxy
8	49	42.2	759	1	SC63_HUMAN Q9ugp8 homo sapien
9	49	42.2	759	1	SC63_MOUSE Q9vhe0 mus musculu
10	47	40.5	418	1	VS12_REOVL P13134 recovirus (c
11	46	39.7	418	1	VS12_REOVL P32922 recovirus (c
12	46	39.7	3060	1	BPEB_HUMAN Q8wxk8 homo sapien
13	45.5	39.2	560	1	PMG2_ARATH Q9m6k1 arabidopsis
14	45	38.8	644	1	SYM_CICAB Q97w55 clostridium
15	45	38.8	657	1	CSP1_CORGL Q01377 coynebacte
16	45	38.8	2558	1	Q9y719 schizosacch
17	44.5	38.4	241	1	HEB1_HABIN P45991 haemophilus
18	44.5	38.4	241	1	HEB2_HABIN P45991 haemophilus
19	44	37.9	226	1	PYRH_SULTO Q971s6 sulfobius
20	44	37.9	569	1	SILF_MOUSE Q921s6 mus musculu
21	44	37.9	837	1	ROD1_YEAST Q02805 saccharomy
22	43.5	37.5	825	1	ILAR_HUMAN P24394 homo sapien
23	43	37.1	267	1	COML_NEIGO Q05085 neisseria g
24	43	37.1	267	1	COML_NEIMA Q9jvb7 neisseria m
25	43	37.1	267	1	COML_NEIMA Q9jvb7 neisseria m
26	43	37.1	418	1	VS12_REOVL P03525 reovirus (c
27	43	37.1	443	1	AM2A_ORYSA P27335 oryza sativ
28	43	37.1	443	1	AM2C_ORYSA P27341 oryza sativ
29	43	37.1	480	1	DISA_TRIGA P15503 trimersura
30	43	37.1	6885	1	SNE2_HUMAN Q8w8h0 homo sapien
31	42.5	36.6	451	1	MYB8_DICDI P34127 dicystoseli
32	42.5	36.6	661	1	PSAB_PROHO P58387 prochloroth
33	42	36.2	405	1	HPCE_ECOLI P37352 e homoprote

34	42	36.2	429	1	HPAG_ECOLI Q46978 e 4-hydroxy
35	42	36.2	447	1	GNT1_RAT Q03325 rattus norv
36	42	36.2	461	1	TRIA_RAT P22934 rattus norv
37	42	36.2	989	1	RPOC_LEUME P44892 leucostoc
38	42	36.2	1097	1	LIFR_HUMAN P43702 homo sapien
39	41.5	35.8	511	1	UI71_CAEL Q20069 caenorhabdi
40	41.5	35.8	1616	1	VAED_SCHPO Q09853 schizosacch
41	41	35.3	305	1	SGAU_MYCPN P75294 mycoplasma
42	41	35.3	438	1	GATD_PYRPU Q8u0x0 pyrococcus
43	41	35.3	576	1	CBPS_YEAST P27614 saccharomy
44	41	35.3	591	1	CALX_MOUSE P35564 mus musculu
45	41	35.3	591	1	CALX_RAT P35565 rattus norv

ALIGNMENTS

RESULT 1

ID GTFD_STRMU STANDARD; PRT; 1462 AA.
AC P49331; O69383; O69386; O69389; O69392; O69398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OK NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme.";
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype B, MT4251 / Serotype F, MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=96231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S., Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / Serotype C, J700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y., Li S., Zhu H., Najjar F., Lai H., White T., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.
CC -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.
CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M29296; AAA26895.1; -
 CC EMBL; D88653; BAA26103.1; -
 CC EMBL; D88656; BAA26107.1; -
 CC EMBL; D88659; BAA26111.1; -
 CC EMBL; D88662; BAA26115.1; -
 CC EMBL; D89979; BAA26121.1; -
 CC EMBL; AE014932; AAN58619.1; -
 CC InterPro; IPR002479; CW_binding.
 CC InterPro; IPR003318; Glyco_hydro_70.
 CC Pfam; PF02324; Glyco_hydro_70; 1.
 CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 KM
 FT SIGNAL 1 ?
 FT CHAIN 1462
 FT DOMAIN 1232 1423
 FT REPEAT 1232 1295
 FT REPEAT 1296 1359
 FT REPEAT 1360 1423
 FT REPEAT 10 10
 FT VARIANT 19
 FT VARIANT 19
 FT VARIANT 58
 FT VARIANT 68
 FT VARIANT 81
 FT VARIANT 113
 FT VARIANT 122
 FT VARIANT 132
 FT VARIANT 135
 FT VARIANT 137
 FT VARIANT 202
 FT VARIANT 255
 FT VARIANT 275
 FT VARIANT 288
 FT VARIANT 301
 FT VARIANT 313
 FT VARIANT 317
 FT VARIANT 328
 FT VARIANT 350
 FT VARIANT 628
 FT VARIANT 688
 FT VARIANT 726
 FT VARIANT 762
 FT VARIANT 762
 FT VARIANT 964
 FT VARIANT 1019
 FT VARIANT 1060
 FT VARIANT 1060
 FT VARIANT 1080
 FT VARIANT 1142
 FT VARIANT 1142
 FT VARIANT 1198
 FT VARIANT 1220
 FT VARIANT 1280
 FT VARIANT 1282
 FT VARIANT 1290
 FT VARIANT 1311

POTENTIAL.
 3 x 63 AA APPROXIMATE TANDEM REPEATS.
 1.
 2.
 3.
 Y -> H (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 I -> V (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 K -> E (IN STRAIN MT4467).
 A -> S (IN STRAINS MT4239 AND MT4245).
 A -> T (IN STRAINS MT4251 AND MT8148).
 T -> I (IN STRAINS MT4239 AND MT4245).
 A -> V (IN STRAINS MT4239, MT4245 AND
 MT8148).
 A -> S (IN STRAINS GS-5 AND MT4467).
 A -> V (IN STRAIN MT4245).
 A -> T (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 V -> L (IN STRAIN MT4239).
 D -> N (IN STRAIN MT8148).
 E -> D (IN STRAINS MT4239, MT4245 AND
 MT4251).
 D -> N (IN STRAINS MT4239, MT4245 AND
 MT4251).
 Q -> H (IN STRAIN MT4245).
 D -> N (IN STRAINS MT4239 AND MT4251).
 E -> K (IN STRAIN MT4239).
 V -> F (IN STRAIN MT4239).
 F -> L (IN STRAINS MT4239, MT4251 AND
 MT4467).
 KKRYT -> EKEYTL (IN STRAIN MT4251).
 A -> S (IN STRAIN MT4239).
 TDGSGA -> ADKGNS (IN STRAIN MT4251).
 TDGGS -> ADKGN (IN STRAINS MT4239 AND
 MT4245).
 T -> A (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 D -> Y (IN STRAIN MT4251).
 E -> K (IN STRAINS MT4245 AND MT4251).
 LG -> IR (IN STRAIN MT4251).
 G -> R (IN STRAIN MT4245).
 G -> R (IN STRAIN MT4239).
 G -> R (IN STRAIN GS-5).
 S -> Q (IN STRAIN MT4239).
 H -> Q (IN STRAIN MT4239).
 Y -> C (IN STRAINS MT4251 AND MT4467).
 F -> L (IN STRAIN MT4245).
 Q -> P (IN STRAIN MT4245).
 K -> T (IN STRAIN MT4245).
 N -> D (IN STRAIN MT4245).

FT VARIANT 1403 1403 G -> D (IN STRAINS GS-5 AND MT4467).
 FT VARIANT 1425 1425 G -> R (IN STRAIN GS-5).
 FT VARIANT 1449 1449 R -> K (IN STRAIN MT4467).
 FT CONFLICT 1428 1462 RYDKNSGNMVYVTLANGRRIGIDRWGIARY -> VY
 R (IN REF. 1).
 SQ SEQUENCE 1462 AA; 163387 MW; CEA4279C4D708645 CRC64;
 Query Match 100.0%; Score 116; DB 1; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AINHLSTLEAMSDNDPOYKND 21
 DB 495 AINHLSTLEAMSDNDPOYKND 515
 RESULT 2
 GFPS_STRDO STANDARD; PRT; 1365 AA.
 ID GFPS_STRDO
 AC P29336;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucosyltransferase-5 precursor (EC 2.4.1.5) (GNF-S) (Dextranucrase)
 DS (Sucrose 6-glucosyltransferase).
 GN GFPS.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OK NCBI_TaxId=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MFE28;
 RX MEDLINE=90316665; PubMed=2142479;
 RA Gilmore K.S., Russell R.R., Ferretti J.J.;
 RT Analysis of the Streptococcus downei gtfS gene, which specifies a
 RT glucosyltransferase that synthesizes soluble glucans.";
 RL Infect. Immun. 58:2452-2458(1990).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
 CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
 CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
 CC PRIMER GLUCAN UNLIKE GTF-1.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
 CC 1,6-GLUCOSE).
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S.MUTANS.
 CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M30943; AAA26898.1; -
 CC InterPro; IPR002479; CW_binding.
 CC InterPro; IPR003318; Glyco_hydro_70.
 CC Pfam; PF01473; CW_binding_1; 8.
 CC Pfam; PF02324; Glyco_hydro_70; 1.
 CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 KW SIGNAL 1
 FT CHAIN 37 1365
 FT REPEAT 157 177
 FT REPEAT 178 197
 FT DOMAIN 198 1061
 FT REPEAT 1062 1082
 OR 37 (POTENTIAL).
 GLUCOSYLTRANSFERASE-S.
 CELL WALL BINDING 1.
 CELL WALL BINDING 2.
 CATALYTIC (APPROXIMATE).
 CELL WALL BINDING 3.


```

FT VARIANT 1402 1402 S -> G (IN STRAINS GS-5, MT4239, MT4467
FT VARIANT 1459 1459 AND MT8148).
FT VARIANT 570 570 Y -> H (IN STRAIN MT4467).
FT CONFLICT 800 817 R -> A (IN REF. 1).
FT CONFLICT 800 817 ADQVRAVASTAPSTDGK -> LIMEFLRLARPHQMA
      (IN REF. 1).
FT CONFLICT 1310 1310 H -> L (IN REF. 1).
FT CONFLICT 1476 AA; 165846 MW; 9C6E09F731B4C6CF CRC64;
SQ SEQUENCE 1476 AA; 165846 MW; 9C6E09F731B4C6CF CRC64;

Query Match 64.7%; Score 75; DB 1; Length 1476;
Best Local Similarity 71.4%; Pred. No. 0.002;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANHSTIEAWSNDNDPOYKND 21
Db 481 ANHSTIEAWSNDNDPTIYHD 501

RESULT 4
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
OS (Sucrose 6-glucosyltransferase).
OC Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D90213; BAAL4241.1;
CC InterPro; IPR002479; CM binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CM binding_1; 13.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
CC SIGNAL 1
CC CHAIN 39 1592 GLUCOSYLTRANSFERASE-1.

```

```

FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MW; BCOA65D079351ECF CRC64;

Query Match 63.8%; Score 74; DB 1; Length 1592;
Best Local Similarity 66.7%; Pred. No. 0.0031;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANHSTIEAWSNDNDPOYKND 21
Db 477 ANHSTIEAWSNDNDPTIYHD 497

RESULT 5
GTF1_STRDO STANDARD; PRT; 1597 AA.
ID GTF1_STRDO
AC P11001;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
OS (Sucrose 6-glucosyltransferase).
OC Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Galpin M.L., Russell R.R.B.;
RA "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RA sobrinus MFE28.";
RL J. Bacteriol. 163:4271-4276(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M17391; AAC63063.1;
CC InterPro; IPR002479; CM binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CM binding_1; 16.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
CC KM

```

	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS
CC	-1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl(1N) = D-fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- DISEASE: DENTAL CARIES.
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCCOSE). GTF-S1 SYNTHESIZES BOTH FORMS OF GLUCANS.
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S.MUTANS.
CC	-1- SIMILARITY: Contains 5 cell wall binding repeats.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.sib.ch).
CC	-----
DR	EMBL; M22054; AAA88592.1; -
DR	EMBL; D88652; BAA26102.1; -
DR	EMBL; D88655; BAA26106.1; -
DR	EMBL; D88658; BAA26110.1; -
DR	EMBL; D88661; BAA26114.1; -
DR	EMBL; D89978; BAA26120.1; -
DR	EMBL; AE014940; AAN58706.1; -
DR	EMBL; M17361; AAA88589.1; -
DR	InterPro; IPR002479; CW_binding.
DR	InterPro; IPR003318; Glyco_hydro_70.
DR	Pfam; PF01473; CW_binding_1; 1.
DR	Pfam; PF02324; Glyco_hydro_70; 1.
KW	Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW	Complete proteome.
FT	SIGNAL 1 34
FT	CHAIN 35 1455
FT	DOMAIN 35 1050
FT	DOMAIN 1126 1455
FT	DOMAIN 1126 1455
FT	REPEAT 1126 1159
FT	REPEAT 1127 1200
FT	REPEAT 1227 1238
FT	REPEAT 1253 1303
FT	REPEAT 1318 1330
FT	REPEAT 1253 1303
FT	REPEAT 1253 1303
FT	VARIANT 21 21
FT	VARIANT 81 81
FT	VARIANT 106 106
FT	VARIANT 116 116
FT	VARIANT 126 126
FT	VARIANT 150 151
FT	VARIANT 256 256
FT	VARIANT 425 425
FT	VARIANT 519 519
FT	VARIANT 538 538
FT	VARIANT 545 545
FT	VARIANT 597 597
FT	VARIANT 600 600
FT	VARIANT 601 601
FT	VARIANT 614 614
FT	VARIANT 727 727
FT	VARIANT 734 734
FT	VARIANT 964 964
FT	VARIANT 1113 1113
FT	VARIANT 1118 1118
FT	VARIANT 1204 1204
FT	GLUCOSYLTRANSFERASE-SI.
FT	CATALYTIC (APPROXIMATE).
FT	GLUCAN-BINDING (APPROXIMATE).
FT	2.4.A, 1 C AND 1 AC REPEATS.
FT	A REPEAT.
FT	C REPEAT.
FT	AC REPEAT.
FT	A REPEAT (INCOMPLETE).
FT	V -> I (IN STRAIN GS-5).
FT	F -> L (IN STRAIN MT4239).
FT	D -> V (IN STRAIN GS-5).
FT	S -> A (IN STRAINS GS-5 AND MT4467).
FT	A -> T (IN STRAIN GS-5).
FT	SR -> PK (IN STRAINS GS-5, MT4239 AND MT4467).
FT	A -> V (IN STRAINS GS-5 AND MT4467).
FT	R -> N (IN STRAIN MT4251).
FT	Y -> D (IN STRAINS MT4245 AND MT4251).
FT	R -> K (IN STRAINS MT4245 AND MT4251).
FT	N -> F (IN STRAINS MT4245 AND MT4251).
FT	N -> D (IN STRAINS MT4245, MT4251, MT4467 AND MT8148).
FT	R -> K (IN STRAINS MT4245, MT4251, MT4467 AND MT8148).
FT	A -> T (IN STRAIN GS-5).
FT	M -> T (IN STRAIN GS-5).
FT	T -> I (IN STRAIN MT8148).
FT	A -> V (IN STRAIN MT8148).
FT	L -> Y (IN STRAIN MT4239).
FT	N -> Y (IN STRAIN MT4239).
FT	A -> T (IN STRAIN MT4239).
FT	I -> V (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).

```

FT 1208 1208 V -> I (IN STRAIN MT8148).
FT 1292 1294 DGH -> NGY (IN STRAINS GS-5, MT4467 AND
FT 1305 1369 MISSING (IN STRAIN MT4245).
FT 1326 1326 I -> V (IN STRAIN MT8148).
FT 1331 1331 T -> A (IN STRAIN GS-5, MT4239, MT4467
FT 1331 1331 AND MT8148).
FT 1377 1377 R -> K (IN STRAIN MT8148).
FT 1398 1398 V -> I (IN STRAIN MT8148).
FT 1424 1424 D -> N (IN STRAIN MT4239).
FT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT 1444 1444 S -> P (IN STRAIN MT8148).
FT 1455 1455 QRLYKSNQVQAKGELLIRGCRICVYDNGNGNENRYVR
FT 1337 1337 TSSGNWYFQNDGALIGMHWYVEGRVYFDENGVRYSHD
FT 1337 1337 QRNHWYDVRDPRGSSAVAFRHSRNGFDNFRF ->
FT 1337 1337 HASILSLMVFRLRBSLSQVKKVSNMTLIPMKFVIWM
FT 1337 1337 (IN REF. 1).
SQ SEQUENCE 1455 AA; 162965 MW; 3CB455A9A4FEC86 CRC64;

Query Match 56.9%; Score 66; DB 1; Length 1455;
Best Local Similarity 66.7%; Pred. No. 0.047;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPOYNKD 21
DB 507 AINHLSTLEAWSNDPOYNKD 527

RESULT 7
HPAG_SALDU STANDARD; PRT; 429 AA.
ID HPAG_SALDU
AC Q9RPU5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 4-hydroxyphenylacetate degradation bifunctional
DE isomerase/decarboxylase [includes: 2-hydroxyhepta-2,4-diene-1,7-dioate
DE isomerase (EC 5.3.3.-) (HHD isomerase); 5-carboxymethyl-2-oxo-hex-3-
DE ene-1,7-dioate decarboxylase (EC 4.1.1.-) (OPE decarboxylase)].
GN HPAG.
OS Salmonella dublin.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=98360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2229;
RA Galinov E.E., Wood M.W., Hedges S.;
RT "Characterization of the hpa genetic locus from Salmonella dublin.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: DECARBOXYLATES OPEPT (5-OXO-PENT-3-ENE-1,2,5-
CC TRICARBOXYLIC ACID) INTO HHD (2-HYDROXY-HEPT-2,4-DIENE-1,7-
CC DIOATE) AND ISOMERIZES IT TO OHD (2-OXO-HEPT-3-ENE-1,7-DIOATE).
CC 1- PATHWAY: 4-hydroxyphenylacetate catabolism.
CC 1- SIMILARITY: Belongs to the FAH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF144422; AAD53501.1; -.
DR InterPro; IPR002529; FAA hydrolyase.
DR Pfam; PF01557; FAA hydrolyase; 2.
KW Aromatic hydrocarbons catabolism; Lyase; Isomerase; Repeat;
KM Multifunctional enzyme.
FT REPEAT 1 215 APPROXIMATE.
FT REPEAT 216 429 APPROXIMATE.
SQ SEQUENCE 429 AA; 47185 MW; D32723999AC4084E CRC64;

```

```

Query Match 44.0%; Score 51; DB 1; Length 429;
Best Local Similarity 47.6%; Pred. No. 2.3;
Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 AINHLSTLEAWSNDPOYN 19
DB 9 AINHLSTLEAWSNDPOYN 29

RESULT 8
SC63_HUMAN STANDARD; PRT; 759 AA.
ID SC63_HUMAN
AC Q9UGP8; O95380; Q9NTE0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Translocation protein Sec63 homolog.
DE SEC63 OR SEC63L.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20008798; PubMed=10543453;
RA Skowronek M.H., Rotter M., Haas I.G.;
RT "Molecular characterization of a novel mammalian DnaJ-like Sec63p
RT homolog.";
RT Biol. Chem. 380:1133-1138(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Hartmann E., Prehn S.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 126-759 FROM N.A.
RA TISSUE=Testis;
RA Poulik A., Klein M., Nemes H.-W., Gassenhuber J., Wiemann S.;
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: Required for preprotein translocation.
CC 1- SUBUNIT: COMPLEX THAT CONTAINS SEC61, SEC62 AND SEC63.
CC 1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC 1- SIMILARITY: Contains 1 J domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF011779; CAB46275.1; -.
DR EMBL; AF100144; AAC83375.1; -.
DR EMBL; AL137338; CAB70701.1; -.
DR PIR; T46504; T46504.
DR HSSP; P25685; 1HDJ.
DR Gene; HGNC:21082; SEC63.
DR InterPro; IPR001623; DnaJ N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR InterPro; IPR004179; Sec63.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF02889; Sec63; 1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; FALSE_NEG.
DR PROSITE; PS50076; DnaJ_2; 1.
KW Chaperone; Transmembrane; Protein transport; Endoplasmic reticulum;
KW Microsome.
FT INIT MET 0
FT DOMAIN 1 13 BY SIMILARITY.
FT TRANSMEM 14 34 LUMENAL (POTENTIAL).
FT DOMAIN 35 68 POTENTIAL.
FT DOMAIN 35 68 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 69 89 POTENTIAL.
FT DOMAIN 90 187 LUMENAL (POTENTIAL).
FT TRANSMEM 188 208 POTENTIAL.
FT DOMAIN 209 759 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 103 164 J-DOMAIN.
FT DOMAIN 733 759 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 759 AA; 87865 MW; BA13AB5B2C53F08 CRC64;

Query Match 42.2% Score 49; DB 1; Length 759;
Best Local Similarity 52.9%; Pred. No. 9.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 5 LSI:EWASNDPQYKND 21
Db 246 IMVLAGASEFDPQYKND 262

RESULT 9
SC63 MOUSE STANDARD; PRT; 759 AA.
AC Q8VHE0; Q8VHE9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Translocation protein SC63 homolog.
CN SEC63 OR SEC63L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Mack M., Noben-Trauth K.;
RT "Phenotypic characterization and positional mapping of the mouse
RT deafness mutation jackson circler (jc).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

[2]
SEQUENCE FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares W.B., Bonaldo W.F., Casavant T.L., Schetz T.R.,
RA Brownstein M.J., Ustin T.B., Toshynski S., Carrinci P., Prange C.,
RA Raha S.S., Loggialano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzyzanski M.J., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Required for preprotein translocation (By similarity).
CC -1- SUBUNIT: COMPLEX THAT CONTAINS SEC61, SEC62 AND SEC63 (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -1- SIMILARITY: Contains 1 J domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AY024346; AAK00580.1; -
CC DR EMBL: BC019366; AAH19366.1; -
CC DR EMBL: BC031846; AAH31846.1; -
CC DR MGI: 2155302; Sec63.
CC DR InterPro: IPR001623; DnaJ_N.
CC DR InterPro: IPR003095; Hsp_DnaJ.
CC DR InterPro: IPR004179; Sec63.
CC DR Pfam: PF00226; DnaJ; 1.
CC DR Pfam: PF02889; Sec63; 1.
CC DR PRINTS: PR00625; DNAJPROTEIN.
CC DR SMART: SM00271; DnaJ; 1.
CC DR SMART: SM00611; SEC63; 1.
CC DR PROSITE: PS00636; DNAJ_1; FALSE_NEG.
CC DR PROSITE: PS50076; DNAJ_2; 1.
CC KW Chapterone; Transmembrane; Protein transport; Endoplasmic reticulum;
CC Microsome.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT DOMAIN 1 13 LUMENAL (POTENTIAL).
CC FT TRANSMEM 14 34 POTENTIAL.
CC FT DOMAIN 35 68 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 69 89 POTENTIAL.
CC FT DOMAIN 90 187 LUMENAL (POTENTIAL).
CC FT TRANSMEM 188 208 POTENTIAL.
CC FT DOMAIN 209 759 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 103 164 J-DOMAIN.
CC FT DOMAIN 733 759 ASP/GLU-RICH (HIGHLY ACIDIC).
CC FT CONFLICT 47 47 N -> S (IN REF. 1).
CC FT CONFLICT 266 266 K -> R (IN REF. 1).
SQ SEQUENCE 759 AA; 87710 MW; D1406C01F07232DA CRC64;

Query Match 42.2% Score 49; DB 1; Length 759;
Best Local Similarity 52.9%; Pred. No. 9.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 5 LSI:EWASNDPQYKND 21
Db 246 IMVLAGASEFDPQYKND 262

RESULT 10
VS12_REV01 STANDARD; PRT; 418 AA.
ID VS12_REV01
AC P11314;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Sigma 2 protein (Core protein).
GN S2.
OS Reovirus (type 1 / strain lang).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_Taxid=10884;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=92015462; PubMed=1920614;
RA Dermody T.S., Schiff L.A., Nibert M.L., Coombs K.M., Fields B.N.;
RT "The S2 gene nucleotide sequences of prototype strains of the three
RT reovirus serotypes: characterization of reovirus core protein sigma
RT 2.";
RT J. Virol. 65:5721-5731(1991).
RN [2]
RP SEQUENCE FROM N.A.
RP MEDLINE=88024195; PubMed=3663211;
RA George C.X., Crowe A., Munemitsu S.M., Atwater J.A., Samuel C.E.;
RT "Biosynthesis of reovirus-specified polypeptides. Molecular cDNA
RT cloning and nucleotide sequence of the reovirus serotype 1 lang
RT strain s2 mRNA which encodes the virion core polypeptide sigma 2.";
RT Biochem. Biophys. Res. Commun. 147:1153-1161(1987).
CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 323
CC ONWARD AND IS SHORTER (135 AA) DUE TO A FRAMESHIFT.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19774; AAA4723.1; -
DR EMBL; L17598; AAA47278.1; ALT_FRAME.
DR PIR; A41306; FOXRL2.
DR InterPro; IPR004317; Sigma_1_2.
DR Pfam; PF03084; Sigma_1_2; 1.
DR ProDom; PD004419; Sigma_1_2; 1.
KW Core protein.
FT CONFLICT 89 89 I -> V (IN REF. 2).
FT CONFLICT 205 205 H -> Y (IN REF. 2).
FT SEQUENCE 418 AA; 47110 MW; 4830CF9217BA1DDB CRC64;
SQ
Query Match 40.5%; Score 47; DB 1; Length 418;
Best Local Similarity 61.5%; Pred. No. 9.4;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 LEAMSDNDPOYXK 20
DB 372 LEAMAREDDOYXK 384

RESULT 11
VS12 REOVJ STANDARD; PRT; 418 AA.
AC P32922;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Sigma 2 protein (Core protein).
OS Reovirus (type 2 / strain D5/Jones).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10885;
RX MEDLINE=92015462; PubMed=1920614;
RA Dermody T.S., Schiff L.A., Nibert K.M., Fields B.N.;
RT "The S2 gene nucleotide sequences of prototype strains of the three
RT reovirus serotypes: characterization of reovirus core protein sigma
RT 2".
RL J. Virol. 65:5721-5731(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19775; AAA47248.1; -
DR PIR; B41306; FOXRJS.
DR InterPro; IPR004317; Sigma_1_2.
DR Pfam; PF03084; Sigma_1_2; 1.
DR ProDom; PD004419; Sigma_1_2; 1.
KW Core protein.
SQ SEQUENCE 418 AA; 47045 MW; E3D4351F55524213 CRC64;
Query Match 39.7%; Score 46; DB 1; Length 418;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 8 LEAMSDNDPOYXK 20
DB 372 LEAMAREDDOYXK 384

RESULT 12

```

```

BPEB_HUMAN
ID BPEB_HUMAN STANDARD; PRT; 3060 AA.
AC Q8WXR8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bullous pemphigoid antigen 1, isoform 7 (Bullous pemphigoid antigen
DE (BPA) (hemidesmosomal plaque protein) (Dystonia musculorum protein)
DE (Fragment).
GN BPA1 OR DMH OR DT OR KIAA0728.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
(1)
SEQUENCE FROM N.A. (ISOFORM 7), AND TISSUE SPECIFICITY.
RP TISSUE=Keratinocytes;
RC MEDLINE=21839111; PubMed=11751855;
RX Okumura M., Yamakawa H., Ohara O., Owaribe K.;
RA "Novel alternative splicings of BPA1 (bullous pemphigoid antigen 1)
RA including the domain structure closely related to MACF (microtubule
RA actin cross-linking factor).".
RL J. Biol. Chem. 277:6682-6687(2002).
(2)
DISEASE.
RP TISSUE=Keratinocytes;
RC MEDLINE=92011493; PubMed=1717441;
RA Sawamura D., Li K., Chu M.-L., Uitto J.;
RA "Human bullous pemphigoid antigen (BPA1). Amino acid sequences
RA deduced from cloned cDNAs predict biologically important peptide
RA segments and protein domains.";
RL J. Biol. Chem. 266:17784-17790(1991).
(1)
FUNCTION: Cytoskeletal linker protein. Anchors keratin-containing
intermediate filaments to the inner plaque of hemidesmosomes. The
proteins may self-aggregate to form filaments or a two-dimensional
mesh (By similarity).
(2)
SUBUNIT: Homodimer (By similarity).
(3)
SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
(4)
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=10;
Name=1; Synonyms=Eb;
Isoid=Q8WXR8-1; Sequence=Displayed;
Name=2;
Isoid=Q03001-1; Sequence=External;
Name=3; Synonyms=1e;
Isoid=Q03001-2; Sequence=External;
Name=4;
Isoid=Q03001-3; Sequence=External;
Name=5;
Isoid=Q03001-4; Sequence=External;
Name=6; Synonyms=EA;
Isoid=Q04833-2; Sequence=External;
Name=7;
Isoid=Q03001-6; Sequence=External;
Name=8;
Isoid=Q03001-5; Sequence=External;
Name=9;
Isoid=Q04833-3; Sequence=External;
Name=10;
Isoid=Q04833-1; Sequence=External;
Note=No experimental confirmation available;
(1)
TISSUE SPECIFICITY: Highly expressed in skeletal muscle and
cultured keratinocytes.
(2)
DISEASE: BPA1 is an autoantigen of bullous pemphigoid
(MIM:600088), an autoimmune subepithelial skin blistering disease.
(3)
SIMILARITY: Belongs to the plakins or cytolinker family.
(4)
SIMILARITY: Contains 4 spectrin repeats.
(5)
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC EMBL, AF000227; AAL62062.1; -

DR MIM; 113810; -

DR MIM; 600088; -

DR GO; GO:0005200; F:structural constituent of cytoskeleton; ISS.

DR GO; GO:0045104; P:intermediate filament cytoskeleton organiza. . .; ISS.

DR InterPro; IPR001101; Plectin_repeat.

DR InterPro; IPR001452; SH3.

DR InterPro; IPR002017; Spectrin.

DR Pfam; PF00681; SH3; 1.

DR Pfam; PF00435; spectrin; 4.

DR Pfam; PF00250; PLEC; 8.

DR SMART; SM00150; SPC; 4.

DR SMART; SM00150; SPC; 4.

DR Antigen; Actin-binding; Coiled coil; Repeat; SH3 domain; KX structural protein; Cytoskeleton; Cell adhesion; Calcium; KM Calcium-binding; Alternative splicing.

FT DOMAIN 126 160 COILED COIL (POTENTIAL).

FT DOMAIN 312 447 COILED COIL (POTENTIAL).

FT DOMAIN 511 546 COILED COIL (POTENTIAL).

FT DOMAIN 564 616 SH3.

FT DOMAIN 706 813 COILED COIL (POTENTIAL).

FT DOMAIN 871 926 COILED COIL (POTENTIAL).

FT DOMAIN 1013 1042 COILED COIL (POTENTIAL).

FT DOMAIN 1087 1139 COILED COIL (POTENTIAL).

FT DOMAIN 1179 1209 COILED COIL (POTENTIAL).

FT REPEAT 264 341 SPECTRIN 1.

FT REPEAT 349 444 SPECTRIN 2.

FT REPEAT 450 550 SPECTRIN 3.

FT REPEAT 935 1031 SPECTRIN 4.

FT REPEAT 1208 1257 PLECTIN 1.

FT REPEAT 1258 1295 PLECTIN 2.

FT REPEAT 1333 1370 PLECTIN 3.

FT REPEAT 1401 1446 PLECTIN 4.

FT REPEAT 1447 1484 PLECTIN 5.

FT REPEAT 1487 1522 PLECTIN 6.

FT REPEAT 1523 1560 PLECTIN 7.

FT REPEAT 1562 1598 PLECTIN 8.

FT NON_TER 3060 3060

SEQ SEQUENCE 3060 AA; 344985 MW; EFB6DF6D72C14E2 CRC64;

Query Match 39.7%; Score 46; DB 1; Length 3060;

Best Local Similarity 47.1%; Pred. No. 1.4e+02;

Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHTSLIAMSNDPQYN 19

DB 2475 SHTSLIASVTDKDPQGN 2491

RESULT 13

PMG2 ARATH STANDARD; PRT; 560 AA.

ID PMG2 ARATH

AC O9M9K1; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable 2,3-bisphosphoglycerate-independent phosphoglycerate mutase 2 (EC 5.4.2.1) (Phosphoglyceromutase 2) (BPG-independent PGAM 2) (PGAM-1 2).

DE AT3G08590 OR F17014.6.

GN Arabidopsis thaliana (Mouse-ear cress).

OS Arabidopsia thaliana; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; core eudicots; Rosidae; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_Taxid=3702;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=cv. Columbia;

RC MEDLINE=21016720; PubMed=1130713;

RA Salanoubat M., Lemcke K., Rieger M., Ansoerg M., Unseid M.,
RA Fartmann B., Valle G., Gloecker H., Perez-Alonso M., Obermaier B.,
RA Delensy M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,
RA De Simone V., Choiane N., Artiguenave F., Robert C., Bottier P.,
RA Wincker P., Cactolico L., Weissbach J., Sarin W., Queller F.,
RA Schaefer M., Moeller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nakamura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Topo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.H., Nordstark G.,
RA Reichelt J., Scharte M., Schoen O., Bargues M., Terol J., Clément J.,
RA Navarro P., Collado C., Perez-Perez A., Otenmaier B., Duchemin D.,
RA Cooke R., Landie M., Berger-Llanco C., Punelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Corté A., Casacuberta E.,
RA Monfort A., Argilou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schopf H., Rüd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Wates A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltcher J., Sellers P., Gill J.E., Feidlyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneke T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana."

RL Nature 408:820-822(2000).

CC -1- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and 3-phosphoglycerate (By similarity).

CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.

CC -1- COFACTOR: Binds 2 manganese ions (By similarity).

CC -1- PATHWAY: Glycolysis.

CC -1- SUBUNIT: Monomer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE BPG-INDEPENDENT PHOSPHOGLYCERATE MUTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC EMBL; AC012562; AAC51361.1; -

DR InterPro; IPR006124; Metalloenzyme.

DR InterPro; IPR005995; Pgm bpd ind.

DR Pfam; PF01676; Metalloenzyme_1.

DR Pfam; PF004429; Pgm bpd ind. 1.

DR TIGRFAMs; TIGR01307; Pgm bpd ind. 1.

KW Isomerase; Glycolysis; Metal-binding; Manganese.

KW ACT_SITE 82 82 PHOSPHOERINE INTERMEDIATE (By similarity).

FT METAL 29 29 MANGANESE 2 (By similarity).

FT METAL 82 82 MANGANESE 2 (By similarity).

FT METAL 431 431 MANGANESE 1 (By similarity).

FT METAL 435 435 MANGANESE 1 (By similarity).

FT METAL 472 472 MANGANESE 2 (By similarity).

FT METAL 473 473 MANGANESE 2 (By similarity).

FT METAL 502 502 MANGANESE 1 (By similarity).

SEQ SEQUENCE 560 AA; 60763 MW; 4D5AC0807AD788B CRC64;

Query Match 39.2%; Score 45.5; DB 1; Length 560;

Best Local Similarity 50.0%; Pred. No. 23;

Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 5 LSLIAMSNDP-QYN 19

DB 25 LITLDGWSBSPDQYN 40

RESULT 14
 ID SYM CLOAB STANDARD; PRT; 644 AA.
 AC 097EM5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine-tRNA ligase)
 DE (MetRS).
 GN METG OR CAC2991.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NC NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=1466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,
 RA Tatsurov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum";
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -1- FUNCTION: Is required not only for elongation of protein synthesis
 but also for the initiation of all mRNA translation through
 initiator tRNA (fMet) aminoacylation (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
 diphosphate + L-methionyl-tRNA(Met).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC Metg subfamily 2A.
 CC -1- SIMILARITY: Contains 1 tRNA-binding domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE007796; AAK80932.1; -
 CC PIR: A97268; A97268.
 CC HAMAP: MF_01228; fused.1.
 CC InterPro: IPR004495; MetG_Cterm.
 CC InterPro: IPR002300; tRNA-synt_1a.
 CC InterPro: IPR001412; tRNA-synt_1.
 CC InterPro: IPR002304; tRNA-synt_mec.
 CC InterPro: IPR002547; tRNA_bind.
 CC Pfam: PF00133; tRNA-synt_1.1.
 CC Pfam: PF01588; tRNA_bind.1.
 CC PIRSF: PIRSF001528; MetRS_dimerising.1.
 CC PRINTS: PR01041; TRNASTYNTMET.
 CC TIGRFAMs: TIGR00398; metG.1.
 CC TIGRFAMs: TIGR00399; metG_C term.1.
 CC PROSITE: PS00178; AA TRNA_LIGASE_1; 1.
 CC PROSITE: PS50886; TRBD.1.
 DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW RNA-binding; tRNA-binding; Metal-binding; Zinc; Complete proteome.
 FT SITE 14 24
 FT SITE 299 303
 FT DOMAIN 542 644
 FT METAL 129 129 TRNA-BINDING.
 FT METAL 132 132 "KMSKS" REGION.
 FT METAL 146 146 ZINC (BY SIMILARITY).
 FT METAL 149 149 ZINC (BY SIMILARITY).
 FT METAL 149 149 ZINC (BY SIMILARITY).
 FT BINDING 302 302 ATP (BY SIMILARITY).
 SQ SEQUENCE 644 AA; 73618 MW; 30035F8B94B60A7A CRC64;

Query Match 38.8%; Score 45; DB 1; Length 644;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 3 NHLSLEAMSDNDPQYNK 20
 Db 234 NYITALGYGSDNDDELXNK 251
 RESULT 15
 ID CSPI_CORGL STANDARD; PRT; 657 AA.
 AC 001377;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE PSI protein precursor.
 DE CSPI OR CGL2875.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 NC NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 17965 / Melassecola;
 RX MEDLINE=93023863; PubMed=1406274;
 RA Joffe G., Mathieu L., Hahn V., Bayan N., Duchiron F., Renaud M.,
 RA Shechter E., Leblon G.;
 RT "Cloning and nucleotide sequence of the cspl gene encoding PSI, one
 of the two major secreted proteins of Corynebacterium glutamicum: the
 deduced N-terminal region of PSI is similar to the Mycobacterium
 RT Mol. Microbiol. 6:2349-2362(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032";
 RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ONE OF THE TWO MAJOR SECRETED PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: THE N-TERMINAL IS VERY SIMILAR TO THE COMPLETE
 SEQUENCES OF THE MYCOBACTERIAL ANTIGENS 85-A, 85-B & 85-C.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X66078; CAA46877.1; -
 CC EMBL: AF005283; BAC00269.1; -
 CC PIR: S25184; S25184.
 CC HSSP: P31953; IDQ2.
 CC InterPro: IPR000801; Esterase_put.
 CC InterPro: IPR000379; Ser_estrs_site.
 CC Pfam: PF00756; Esterase.1.
 DR Signal; Complete proteome; Pyrrolidone carboxylic acid.
 KW SIGNAL
 FT CHAIN 44 657
 FT MOD RES 44 44
 FT CONFLICT 54 54 I -> V (IN REF. 1).
 FT CONFLICT 177 177 V -> I (IN REF. 1).
 FT CONFLICT 189 189 G -> E (IN REF. 1).
 FT CONFLICT 320 320 G -> T (IN REF. 1).
 FT CONFLICT 330 330 A -> T (IN REF. 1).
 FT CONFLICT 350 350 S -> N (IN REF. 1).
 FT CONFLICT 461 461 A -> D (IN REF. 1).
 FT CONFLICT 562 562 E -> D (IN REF. 1).
 FT CONFLICT 596 596 A -> V (IN REF. 1).
 FT CONFLICT

SQ SEQUENCE 657 AA; 70674 MW; E2C3F1B57A7938A3 CRC64;
 Query Match 38.8%; Score 45; DB 1; Length 657;
 Best Local Similarity 53.8%; Pred. No. 33;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 9 EAMSDNDPQYNKD 21
 Db 286 ERWQENDPKSNVD 298

Search completed: November 13, 2003, 09:45:30
 Job time : 8.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-11
Sequence: 1 AINHLSTLEAWSDNDPQYKND 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_ricent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	90.5	1575	2 Q9LCH3	Q9LCH3 streptococc
2	103	88.8	1577	2 Q54178	Q54178 streptococc
3	90	77.6	1577	2 Q55265	Q55265 streptococc
4	85	73.3	1449	2 Q68542	Q68542 streptococc
5	85	73.3	1449	2 Q55264	Q55264 streptococc
6	81	69.8	1338	2 Q9XWJ4	Q9XWJ4 streptococc
7	81	69.8	1554	2 Q8XJL5	Q8XJL5 streptococc
8	78	67.2	1477	2 Q9L466	Q9L466 leuconostoc
9	78	67.2	1508	2 Q9EZH5	Q9EZH5 leuconostoc
10	78	67.2	1508	2 Q52224	Q52224 leuconostoc
11	76	65.5	1512	2 Q9XWJ5	Q9XWJ5 streptococc
12	74	63.8	1590	2 Q59983	Q59983 streptococc
13	74	63.8	1590	2 Q55263	Q55263 streptococc
14	73	62.9	1016	2 Q9LCT7	Q9LCT7 leuconostoc
15	68	58.6	1518	2 Q00600	Q00600 streptococc
16	68	58.6	1527	2 Q9ZAR4	Q9ZAR4 leuconostoc

17	68	58.6	1527	2 Q8KRE1	Q8KRE1 leuconostoc
18	67	57.8	1539	2 Q00539	Q00539 streptococc
19	64	55.2	2057	2 Q9RE05	Q9RE05 leuconostoc
20	63	54.3	2835	2 Q8G9Q2	Q8G9Q2 leuconostoc
21	51	44.0	429	16 Q827Q2	Q827Q2 salmonella
22	50	43.1	429	16 Q8ZQ49	Q8ZQ49 salmonella
23	50	43.1	474	10 Q9SGV0	Q9SGV0 arabidopsis
24	49	42.2	286	16 Q9RDG2	Q9RDG2 streptomyces
25	49	42.2	470	16 Q8G3R9	Q8G3R9 bifidobacte
26	49	42.2	516	11 Q8XJL5	Q8XJL5 mus musculu
27	49	42.2	536	4 Q8IWL0	Q8IWL0 homo sapien
28	49	42.2	833	2 Q8GN89	Q8GN89 pseudomonas
29	49	42.2	1215	2 Q8GQX4	Q8GQX4 selomonas
30	48.5	41.8	90	2 Q8KZ38	Q8KZ38 uncultured
31	48.5	41.8	448	2 Q8VLPS	Q8VLPS yersinia ru
32	47.5	40.9	369	16 Q8YL13	Q8YL13 anabaena sp
33	47	40.5	175	10 Q9LWJ8	Q9LWJ8 arabidopsis
34	47	40.5	195	10 Q8LDM8	Q8LDM8 arabidopsis
35	47	40.5	403	16 P96402	P96402 mycobacteri
36	47	40.5	418	12 Q85682	Q85682 reovirus sp
37	47	40.5	418	12 Q85681	Q85681 reovirus sp
38	47	40.5	418	12 Q85676	Q85676 reovirus sp
39	47	40.5	418	12 Q85674	Q85674 reovirus sp
40	47	40.5	418	12 Q85675	Q85675 reovirus sp
41	47	40.5	418	12 Q85677	Q85677 reovirus sp
42	47	40.5	418	12 Q85678	Q85678 reovirus sp
43	47	40.5	418	12 Q8VSE2	Q8VSE2 ndelle viru
44	47	40.5	445	16 Q9RTN4	Q9RTN4 deinococcus
45	47	40.5	490	2 Q07382	Q07382 lactococcus

ALIGNMENTS

RESULT 1

ID Q9LCH3 PRELIMINARY; PRT; 1575 AA.
AC Q9LCH3.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DR 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTPF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
encoding glucosyltransferase from Streptococcus oralis.";
RI Infect. Immun. 68:2475-2483 (2000).
DR EMBL; AB025228; BA95201.1; -;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26EAD7C2E543 CRC64;

Query Match 90.5%; Score 105; DB 2; Length 1575;
Best Local Similarity 90.5%; Pred. No. 1.2e-07;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSDNDPQYKND 21
|||
546 AINHLSTLEAWSDNDPQYKND 566

RESULT 2

```

054178
ID 054178 PRELIMINARY; PRT; 1577 AA.
AC 054178; 054247;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, t9g, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC34483.1; -.
DR EMBL; M89776; AAC26969.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW TRANSFERASE.
SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match 88.8%; Score 103; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 2.5e-07;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKND 21
Db 548 ALKHLSTLEAWSNDPQYKND 568

RESULT 3
055265 PRELIMINARY; PRT; 1577 AA.
ID 055265
AC 055265;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTFM.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122197; PubMed=7822030;
RX Simpson C.L., Giffard P.M., Jacques N.A.;
RA "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL; L35928; AAC41413.1; -.
DR InterPro; IPR004829; Csaurface_antigen.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF02324; Glyco_hydro_70; 1.

```

```

DR ProDom; PD153432; Csaurface_antigen; 1.
KW Signal; Transferase. 38
FT SIGNAL 1 POTENTIAL.
FT CHAIN 39 1577 GLUCOSYLTRANSFERASE.
SQ SEQUENCE 1577 AA; 175290 MW; 3EEFB898A7D3A7BF3 CRC64;

Query Match 77.6%; Score 90; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 3e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKND 21
Db 588 AINHLSTLEAWSNDPQYKND 608

RESULT 4
068542 PRELIMINARY; PRT; 1449 AA.
ID 068542
AC 068542;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Taffe R.I.;
RT "Streptococcus salivarius V1477 gtfN";
RT Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW TRANSFERASE.
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BF8 CRC64;

Query Match 73.3%; Score 85; DB 2; Length 1449;
Best Local Similarity 76.2%; Pred. No. 0.00017;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKND 21
Db 536 AINHLSTLEAWSNDPQYKND 556

RESULT 5
055264 PRELIMINARY; PRT; 1449 AA.
ID 055264
AC 055264;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTF.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122197; PubMed=7822030;
RX Simpson C.L., Giffard P.M., Jacques N.A.;
RA "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL; L35495; AAC41412.1; -.

```

DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1.8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MW; DD62F0730686A46 CRC64;

Query Match 73.3%; Score 85; DB 2; Length 1449;
 Best Local Similarity 76.2%; Pred. No. 0.00017;
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYKND 21
 DB 536 AINHLSTLEAWSNDNDPOYKND 556

RESULT 6

Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
 AC Q9WXJ4; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-S.
 GN Streptococcus criceti.
 OS Streptococcus criceti.
 OC Plasmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1333;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=H-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S.criticus glucosyltransferase (gfts and gftf) genes";
 RL Submitted (Mar-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB026123; BAA77236.1;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1.10.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Plasmid.
 SQ SEQUENCE 1338 AA; 146558 MW; 0A30C6E10E15D99B CRC64;

Query Match 69.8%; Score 81; DB 2; Length 1338;
 Best Local Similarity 76.2%; Pred. No. 0.00066;
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYKND 21
 DB 437 AINHLSTLEAWSNDNDPOYKND 457

RESULT 7

Q8KZL5 PRELIMINARY; PRT; 1554 AA.
 AC Q8KZL5; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sanpuku H., Hayakawa M.,
 Mukasa H., Shiroza T., Abiko Y.;

RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan";
 RL Biochim. Biophys. Acta 1570:75-79 (2002).
 DR EMBL; AB089438; BAC07265.1;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 6.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferase.
 SQ SEQUENCE 1554 AA; 171676 MW; 6981BC1DAE24A73 CRC64;

Query Match 69.8%; Score 81; DB 2; Length 1554;
 Best Local Similarity 71.4%; Pred. No. 0.00078;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYKND 21
 DB 485 AINHLSTLEAWSNDNDPOYKND 505

RESULT 8

Q9L466 PRELIMINARY; PRT; 1477 AA.
 AC Q9L466; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRG.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Remaud-Simeon M., Pizut S., Sarracal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternanucrase, a sucrose
 RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355";
 RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ250172; CAB76565.1;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFCB831 CRC64;

Query Match 67.2%; Score 78; DB 2; Length 1477;
 Best Local Similarity 71.4%; Pred. No. 0.0022;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYKND 21
 DB 532 AINHLSTLEAWSNDNDPOYKND 552

RESULT 9

Q9EZHS PRELIMINARY; PRT; 1508 AA.
 AC Q9EZHS; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase Derb742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=B-742CB;

RA Kim H.-S., Kim D., Ryu H.-J., Roby J.F.
 RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
 RI Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF294469; ANG38021.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87AE4F3A CRC64;

Query Match 67.2%; Score 78; DB 2; Length 1508;
 Best Local Similarity 71.4%; Pred. No. 0.0023;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 563 ANOHLSTLEAWSNDNDPEYKXD 583

RESULT 10

OS 052224 PRELIMINARY; PRT; 1508 AA.
 AC 052224;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OC NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
 RT "Cloning and sequencing of a gene coding for an extracellular
 dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
 RT extralisting only a a(1-6) glucan."
 RL FEMS Microbiol. Lett. 0:0-0(1998).
 DR EMBL: AF030129; AAB95453.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 168511 MW; E70CEGB57A70D1F0 CRC64;

Query Match 67.2%; Score 78; DB 2; Length 1508;
 Best Local Similarity 71.4%; Pred. No. 0.0023;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 563 ANOHLSTLEAWSNDNDPEYKXD 583

RESULT 11

OS 09WXJ5 PRELIMINARY; PRT; 1512 AA.
 AC 09WXJ5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF.
 OS Streptococcus criceti.
 OC Bacteriia; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;

RA Inoue M., Fukui K., Miyagi A.
 RT "S.cricetus glucosyltransferase (gtfs and gtfI) genes."
 RI Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB026123; BAA77237.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM Plasmid.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;

Query Match 65.5%; Score 76; DB 2; Length 1512;
 Best Local Similarity 66.7%; Pred. No. 0.0047;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 485 ALAHSTLEAWSNDNYNND 505

RESULT 12

OS 059983 PRELIMINARY; PRT; 1590 AA.
 AC 059983;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase-1 precursor (EC 2.4.1.5).
 GN GTFI.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=QM2176;
 RA MEDLINE-94146405; PubMed-8312602;
 RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
 RT "DNA sequence of the glucosyltransferase gene of serotype d
 RT Streptococcus sobrinus."
 RL DNA Seq. 4:19-27(1993).
 DR EMBL: D13858; BAA02976.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 16.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM Glycosyltransferase; Signal; Transferase.
 FT SIGNAL.
 FT CHAIN 1 38
 FT SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;
 SQ

Query Match 63.8%; Score 74; DB 2; Length 1590;
 Best Local Similarity 66.7%; Pred. No. 0.01;
 Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 477 ANNHVSIVELAWSNDNDPTYLHD 497

RESULT 13

OS 055263 PRELIMINARY; PRT; 1590 AA.
 AC 055263;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JUN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-I.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1310;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase
   produced from Streptococcus sobrinus ATCC 33478."
RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL; D63570; BAA09792.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 15.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match
Best Local Similarity 66.7%; Pred. No. 0.01;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
DB 477 ANNHVSIVEAWSNDNDPTLHD 497

RESULT 14
Q9LCU7 PRELIMINARY; PRT; 1016 AA.
ID Q9LCU7;
AC Q9LCU7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
   mesenteroides NRRL B-512F."
RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
DR EMBL; AB020020; BAA90527.1; -.
DR HSSP; P06278; 1VJ5.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 88966FDE13CCCB47 CRC64;

Query Match
Best Local Similarity 72.2%; Pred. No. 0.0091;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HUSILFAWSNDNDPQYKND 21
DB 551 HVSILDEWSDNDNAEYKCD 568

RESULT 15
Q00600 PRELIMINARY; PRT; 1518 AA.
AC Q00600;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase I (EC 2.4.1.5) (GTF) (Dextranucrase) (Sucrose 6-
   glucosyltransferase).
GN GTFV.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25975;
RX MEDLINE=92148377; PubMed=1838391;
RA Giffard P.M., Simpson C.L., Milward C.F., Jacques N.A.;
RT "Molecular characterization of a cluster of at least two-
   glucosyltransferase genes in Streptococcus salivarius ATCC 25975."
RL J. Gen. Microbiol. 137:2577-2593(1991).
CC CC
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- DISEASE: DENTAL CARIES.
CC EMBL; 211873; CAAT7900.1; -.
CC EMBL; M64111; AAA26896.1; -.
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Repeat; Dental caries.
FT DOMAIN 1307 1482 6 DIRECT REPEATS.
FT REPEAT 1307 1338 REPEAT 1.
FT REPEAT 1339 1352 REPEAT 2.
FT REPEAT 1372 1403 REPEAT 3.
FT REPEAT 1404 1417 REPEAT 4.
FT REPEAT 1437 1468 REPEAT 5.
FT REPEAT 1469 1482 REPEAT 6.
SQ SEQUENCE 1518 AA; 167730 MW; DAA41F717098B59A CRC64;

Query Match
Best Local Similarity 58.6%; Score 68; DB 2; Length 1518;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 19
DB 501 ALAHISVLEAWSLNDNHYN 519

```

Search completed: November 13, 2003, 09:44:03
 Job time : 33.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)

65.165 Million cell updates/sec

Title: US-09-290-049A-11

Perfect score: 116

Sequence: 1 AINHLSTLEAWSNDPQYNKD 21

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	545	US-09-604-957-4	Sequence 4, Appli
2	116	100.0	1430	US-09-008-172-2	Sequence 2, Appli
3	116	100.0	1430	US-09-210-361-6	Sequence 2, Appli
4	116	100.0	1430	US-09-740-274-6	Sequence 6, Appli
5	90	77.6	1577	US-08-793-824-2	Sequence 2, Appli
6	75	64.7	1475	US-09-007-999-2	Sequence 2, Appli
7	75	64.7	1475	US-09-210-361-2	Sequence 2, Appli
8	75	64.7	1475	US-09-740-274-2	Sequence 2, Appli
9	68	58.6	523	US-09-604-957-5	Sequence 5, Appli
10	66	56.9	1375	US-09-210-361-4	Sequence 4, Appli
11	66	56.9	1375	US-09-740-274-4	Sequence 4, Appli
12	64	55.2	584	US-09-604-957-6	Sequence 6, Appli
13	64	55.2	2057	US-09-499-203-2	Sequence 7, Appli
14	53.5	46.1	535	US-09-604-957-7	Sequence 3, Appli
15	53.5	46.1	1278	US-09-604-957-3	Sequence 3, Appli
16	49	42.2	623	US-09-205-258-482	Sequence 482, App
17	49	42.2	704	US-09-205-258-422	Sequence 422, App
18	49	40.5	521	US-09-107-532A-6431	Sequence 6431, Ap
19	46.5	40.1	327	US-09-107-532A-6181	Sequence 6181, Ap
20	46	39.7	31	US-09-308-345A-42	Sequence 42, Appli
21	46	39.7	703	US-09-422-869-27	Sequence 27, Appli
22	46	39.7	703	US-09-308-345A-49	Sequence 49, Appli
23	46	39.7	10182	US-09-134-001C-3159	Sequence 3159, Ap
24	45	38.8	151	US-09-198-452A-815	Sequence 815, App
25	45	38.8	324	US-08-508-761B-31	Sequence 31, Appli
26	45	38.8	657	US-08-508-761B-2	Sequence 2, Appli
27	44.5	38.4	259	US-08-277-231A-3	Sequence 3, Appli

ALIGNMENTS

28	44.5	38.4	259	2	US-08-473-750-6	Sequence 6, Appli
29	44.5	38.4	259	2	US-08-477-326-6	Sequence 6, Appli
30	44	37.9	467	4	US-09-134-001C-4200	Sequence 4200, Ap
31	43.5	37.5	197	3	US-08-897-020-7	Sequence 7, Appli
32	43.5	37.5	197	4	US-09-350-823-7	Sequence 20, Appli
33	43.5	37.5	691	4	US-09-313-942-20	Sequence 18, Appli
34	43.5	37.5	694	4	US-09-313-942-18	Sequence 22, Appli
35	43.5	37.5	694	4	US-09-313-942-22	Sequence 30, Appli
36	43.5	37.5	784	4	US-09-313-942-30	Sequence 32, Appli
37	43.5	37.5	793	4	US-09-313-942-32	Sequence 6, Appli
38	42	36.2	205	3	US-08-684-024-6	Sequence 6, Appli
39	42	36.2	205	3	US-09-145-868-6	Sequence 6, Appli
40	42	36.2	1001	1	US-07-797-556-6	Sequence 2, Appli
41	42	36.2	1001	1	US-07-943-843-2	Sequence 2, Appli
42	42	36.2	1001	1	US-08-347-003-2	Sequence 6, Appli
43	42	36.2	1097	1	US-07-943-843-6	Sequence 6, Appli
44	42	36.2	1097	1	US-08-347-003-6	Sequence 25, Appli
45	41.5	35.8	720	3	US-09-296-284-25	

RESULT 1
US-09-604-957-4
Sequence 4, Application US/09604957

Patent No. 6486314

GENERAL INFORMATION:
APPLICANT: VAN GEL-SCHUTTEN, GERRITJINA HENDRIKA

Query Match 100.0%; Score 116; DB 4; Length 545;
Best Local Similarity 100.0%; Pred. No. 4.3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AINHLSTLEAWSNDPQYNKD 21
Db 75 AINHLSTLEAWSNDPQYNKD 95

RESULT 2
US-09-008-172-2
Sequence 2, Application US/09008172

Patent No. 6127602

GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.

```

; ORGANISM: Streptococcus mutans
US-09-008-172-2
Query Match          100.0%; Score 116; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPQYKND 21
Db 495 AINHLSTLEAMSDNDPQYKND 515

RESULT 3
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match          100.0%; Score 116; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPQYKND 21
Db 495 AINHLSTLEAMSDNDPQYKND 515

RESULT 4
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; EARLIER FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6
```

```

; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match          100.0%; Score 116; DB 4; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPQYKND 21
Db 495 AINHLSTLEAMSDNDPQYKND 515

RESULT 5
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 9957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match          77.6%; Score 90; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 3.1e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPQYKND 21
Db 588 AINHLSTLEAMSYNDHQYKND 608
```

```
RESULT 6
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match
Best Local Similarity 64.7%; Score 75; DB 3; Length 1475;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 481 ANDHSLIEAWSNDNDTPYLHD 501

RESULT 7
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match
Best Local Similarity 64.7%; Score 75; DB 3; Length 1475;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 481 ANDHSLIEAWSNDNDTPYLHD 501

RESULT 8
US-09-740-274-2
```

```
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match
Best Local Similarity 64.7%; Score 75; DB 4; Length 1475;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 481 ANDHSLIEAWSNDNDTPYLHD 501

RESULT 9
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDIJNA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHROUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match
Best Local Similarity 58.6%; Score 68; DB 4; Length 523;
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 75 ANOHLSDIEAWSNDNDPLVYTD 95

RESULT 10
US-09-210-361-4
```



```
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4
```

```
Query Match      56.9%; Score 66; DB 3; Length 1375;
Best Local Similarity 66.7%; Pred. No. 0.025;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOYMD 21
DB 507 ANKHLSTLEAMSYNDPTPLHD 527
```

```
RESULT 11
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4
```

```
Query Match      56.9%; Score 66; DB 4; Length 1375;
Best Local Similarity 66.7%; Pred. No. 0.025;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOYMD 21
DB 507 ANKHLSTLEAMSYNDPTPLHD 527
```

```
RESULT 12
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHROUT, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6
```

```
Query Match      55.2%; Score 64; DB 4; Length 584;
Best Local Similarity 66.7%; Pred. No. 0.019;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOY 18
DB 75 ANKHLSTLEDMWNGKDPY 92
```

```
RESULT 13
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: KUNZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; PRIOR FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2
```

```
Query Match      55.2%; Score 64; DB 4; Length 2057;
Best Local Similarity 66.7%; Pred. No. 0.086;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOY 18
DB 665 ANKHLSTLEDMWNGKDPY 682
```

```
RESULT 14
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
```

```

; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

```

Query Match          46.1%; Score 53.5; DB 4; Length 535;
Best Local Similarity 55.6%; Pred. No. 0.94;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

```

```

QY      4 HSLIEMWSNDPOY-NK 20
        ||:|||||:|:|:|
DB      77 HINILEDWMHADPEYFNK 94

```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

```

```

Query Match          46.1%; Score 53.5; DB 4; Length 1278;
Best Local Similarity 55.6%; Pred. No. 2.7;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

```

```

QY      4 HSLIEMWSNDPOY-NK 20
        ||:|||||:|:|:|
DB      553 HINILEDWMHADPEYFNK 570

```

Search completed: November 13, 2003, 09:47:59
Job time : 13.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-11
Perfect score: 116
Sequence: 1 AINHLSTLEAWSNDNDPQYNKD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues
Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCITIS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	100.0	545	10 US-09-995-749A-10	Sequence 10, Appl
2	116	100.0	1430	9 US-09-740-274-6	Sequence 6, Appl
3	75	64.7	1475	9 US-09-740-274-2	Sequence 2, Appl
4	68	58.6	522	10 US-09-995-749A-11	Sequence 11, Appl
5	66	56.9	1375	9 US-09-740-274-4	Sequence 4, Appl
6	64	55.2	584	10 US-09-995-749A-12	Sequence 12, Appl
7	53.5	46.1	535	10 US-09-995-749A-13	Sequence 13, Appl
8	53.5	46.1	1781	10 US-09-995-749A-2	Sequence 2, Appl
9	51	44.0	429	9 US-09-815-242-14022	Sequence 14022, A
10	49	42.2	268	15 US-10-156-761-12823	Sequence 12823, A
11	49	42.2	623	12 US-09-933-767-482	Sequence 482, App
12	49	42.2	623	15 US-10-023-282-482	Sequence 482, App
13	49	42.2	704	12 US-09-933-767-482	Sequence 482, App
14	49	42.2	704	15 US-10-023-282-482	Sequence 482, App
15	48	41.4	216	15 US-10-156-761-9158	Sequence 9158, Ap

16	47	40.5	95	9 US-09-945-301-14	Sequence 14, Appl
17	46	39.7	647	15 US-10-116-519-10	Sequence 10, Appl
18	46	39.7	703	10 US-09-768-877-27	Sequence 27, Appl
19	45	38.8	657	10 US-09-738-626-6670	Sequence 6670, Ap
20	44.5	38.4	114	15 US-10-156-761-12167	Sequence 12167, A
21	44.5	38.4	1056	12 US-10-098-871-65	Sequence 65, Appl
22	44	37.9	212	15 US-10-156-761-7969	Sequence 7969, Ap
23	44	37.9	300	9 US-09-794-960-5	Sequence 5, Appl
24	43.5	37.5	91	12 US-10-098-871-75	Sequence 75, Appl
25	43.5	37.5	192	10 US-09-796-692-1790	Sequence 1790, Ap
26	43.5	37.5	192	10 US-09-796-692-2054	Sequence 2054, Ap
27	43.5	37.5	192	15 US-10-040-862-1790	Sequence 1790, Ap
28	43.5	37.5	192	15 US-10-040-862-2054	Sequence 2054, Ap
29	43.5	37.5	691	10 US-09-935-868-20	Sequence 20, Appl
30	43.5	37.5	691	15 US-10-282-162-20	Sequence 20, Appl
31	43.5	37.5	691	12 US-10-282-162-20	Sequence 20, Appl
32	43.5	37.5	694	10 US-09-935-868-18	Sequence 18, Appl
33	43.5	37.5	694	10 US-09-935-868-22	Sequence 22, Appl
34	43.5	37.5	694	12 US-10-282-162-18	Sequence 18, Appl
35	43.5	37.5	694	12 US-10-282-162-22	Sequence 22, Appl
36	43.5	37.5	694	15 US-10-287-035-18	Sequence 18, Appl
37	43.5	37.5	694	15 US-10-287-035-22	Sequence 22, Appl
38	43.5	37.5	735	12 US-10-098-871-12	Sequence 12, Appl
39	43.5	37.5	739	12 US-10-098-871-61	Sequence 61, Appl
40	43.5	37.5	739	12 US-10-098-871-62	Sequence 62, Appl
41	43.5	37.5	776	10 US-09-935-868-36	Sequence 36, Appl
42	43.5	37.5	776	10 US-09-935-868-40	Sequence 40, Appl
43	43.5	37.5	776	10 US-09-935-868-44	Sequence 44, Appl
44	43.5	37.5	776	15 US-10-287-035-36	Sequence 36, Appl
45	43.5	37.5	776	15 US-10-287-035-40	Sequence 40, Appl

ALIGNMENTS

```
RESULT 1
US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995, 749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-995-749A-10

Query Match          100.0%; Score 116; DB 10; Length 545;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AINHLSTLEAWSNDNDPQYNKD 21
      |||||
DB      75 AINHLSTLEAWSNDNDPQYNKD 95

RESULT 2
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
```

```
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match      100.0%; Score 116; DB 9; Length 1430;
Best Local Similarity 100.0%; Pred. No. 4,5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AINHLSTLEAWSNDNDPOYNKD 21
DB      495 AINHLSTLEAWSNDNDPOYNKD 515

RESULT 3
US-09-740-274-2
Sequence 2, Application US/09740274
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match      64.7%; Score 75; DB 9; Length 1475;
Best Local Similarity 71.4%; Pred. No. 0.0091;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 AINHLSTLEAWSNDNDPOYNKD 21
DB      481 AINHLSTLEAWSNDNDPTPLHD 501

RESULT 4
US-09-995-749a-11
Sequence 11, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEL-SCHUTTEN, GERITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 522
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-995-749a-11

Query Match      58.6%; Score 68; DB 10; Length 522;
Best Local Similarity 66.7%; Pred. No. 0.035;
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 AINHLSTLEAWSNDNDPOYNKD 21
DB      75 ANQHLSTLEAWSNDNDPLVYTD 95

RESULT 5
US-09-740-274-4
Sequence 4, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match      56.9%; Score 66; DB 9; Length 1375;
Best Local Similarity 66.7%; Pred. No. 0.2;
```

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AINHLSLEAWSNDNDPOYKD 21
DB 507 ANHLSLEAWSNDNDPYLHD 527

RESULT 6
US-09-995-749A-12

/ Sequence 12, Application US/09995749A
/ Patent No. US20020155568A1
/ GENERAL INFORMATION:
/ APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
/ APPLICANT: DIKHUIZEN, LUBBERT
/ APPLICANT: RAHAOUTI, HAKIM
/ APPLICANT: LEER, ROBERT-JAN
/ TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
/ FILE REFERENCE: BO43388-CIP
/ CURRENT APPLICATION NUMBER: US/09/995,749A
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: 09/604,957
/ PRIOR FILING DATE: 2000-06-28
/ PRIOR APPLICATION NUMBER: EPO 00201871.1
/ PRIOR FILING DATE: 2000-05-25
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 12
/ LENGTH: 584
/ TYPE: PRT
/ ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 55.2%; Score 64; DB 10; Length 584;
Best Local Similarity 66.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSLEAWSNDNDPOY 18
DB 75 ANKHLSLEEDWNGKMPQY 92

RESULT 7
US-09-995-749A-13
/ Sequence 13, Application US/09995749A
/ Patent No. US20020155568A1
/ GENERAL INFORMATION:
/ APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
/ APPLICANT: DIKHUIZEN, LUBBERT
/ APPLICANT: RAHAOUTI, HAKIM
/ APPLICANT: LEER, ROBERT-JAN
/ TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
/ FILE REFERENCE: BO43388-CIP
/ CURRENT APPLICATION NUMBER: US/09/995,749A
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: 09/604,957
/ PRIOR FILING DATE: 2000-06-28
/ PRIOR APPLICATION NUMBER: EPO 00201871.1
/ PRIOR FILING DATE: 2000-05-25
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 13
/ LENGTH: 535
/ TYPE: PRT
/ ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 46.1%; Score 53.5; DB 10; Length 535;
Best Local Similarity 55.6%; Pred. No. 6;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 4 HSLSLAWSNDNDPOY-NK 20
DB 77 HINILEDWNGHADPEYFNK 94

RESULT 8
US-09-995-749A-2

/ Sequence 2, Application US/09995749A
/ Patent No. US20020155568A1
/ GENERAL INFORMATION:
/ APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
/ APPLICANT: DIKHUIZEN, LUBBERT
/ APPLICANT: RAHAOUTI, HAKIM
/ APPLICANT: LEER, ROBERT-JAN
/ TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
/ FILE REFERENCE: BO43388-CIP
/ CURRENT APPLICATION NUMBER: US/09/995,749A
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: 09/604,957
/ PRIOR FILING DATE: 2000-06-28
/ PRIOR APPLICATION NUMBER: EPO 00201871.1
/ PRIOR FILING DATE: 2000-05-25
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 1781
/ TYPE: PRT
/ ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 46.1%; Score 53.5; DB 10; Length 1781;
Best Local Similarity 55.6%; Pred. No. 22;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 4 HSLSLAWSNDNDPOY-NK 20
DB 1056 HINILEDWNGHADPEYFNK 1073

RESULT 9
US-09-815-242-14022
/ Sequence 14022, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14022
/ LENGTH: 429
/ TYPE: PRT

ORGANISM: Salmonella typhi
US-09-815-242-14022

Query Match 44.0%; Score 51; DB 9; Length 429;
Best Local Similarity 47.6%; Pred. No. 11;
Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 AINHLSTLEAMSND--NDPOYN 19
DB 9 AINHLSTLEAMSND--NDPOYN 29

RESULT 10
US-10-156-761-12823
Sequence 12823, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12823
LENGTH: 268
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-12823

Query Match 42.2%; Score 49; DB 15; Length 268;
Best Local Similarity 53.3%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 INHLSTLEAMSNDP 16
DB 220 ICKLSTVLKQWNEDEP 234

RESULT 11
US-09-933-767-482
Sequence 482, Application US/09933767
Publication No. US20030181692A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P2
CURRENT APPLICATION NUMBER: US/09/933,767
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: PCT/US01/05614
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/184,836
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/193,170
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 09/205,258
PRIOR FILING DATE: 1998-12-04
PRIOR APPLICATION NUMBER: PCT/US98/11422
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/048,885
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,375
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,881

PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,880
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,896
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,020
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,876
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,895
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,884
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,894
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,971
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,964
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,882
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,899
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,893
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,900
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,901
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,892
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,915
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,019
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,970
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,972
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,916
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,373
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,875
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,374
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,917
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,949
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,974
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,883
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,897
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,898
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,962
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,963
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,877
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,878
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18

```

; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 482
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (575)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-933-767-482

Query March 42.2% Score 49; DB 12; Length 623;
Best Local Similarity 52.9%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSLLEWMSDNDPOYNKD 21
Db 107 IMVLXGASEPDPQYNKD 123

RESULT 12
; Sequence 482, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; EARLIER FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 482
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
```

```
LOCATION: (111)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (575)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-482

Query Match      42.2%; Score 49; DB 15; Length 623;
Best Local Similarity 52.9%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy      5  LSLFAMSDNDPOXNKD 21
Db      107  IMVLXGASEPDPQNKD 123

RESULT 13
US-09-933-767-422
Sequence 422, Application US/09933767
Publication No. US20030181692A1
GENERAL INFORMATION:
APPLICANT: Nt et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P2
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: US/09/933,767
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/184,836
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/193,170
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 09/205,258
PRIOR FILING DATE: 1998-12-04
PRIOR APPLICATION NUMBER: PCT/US98/11422
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/048,885
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,375
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,881
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,880
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,876
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,895
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,884
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,894
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,971
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,964
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,882
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,899
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,893
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,900
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,901
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,892
PRIOR FILING DATE: 1997-06-06
```

```
PRIOR APPLICATION NUMBER: 60/048,915
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,019
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,970
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,972
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,916
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,373
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,875
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,374
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,917
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,949
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,974
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,883
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,897
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,898
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,962
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,963
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,877
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,878
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/073,160
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,159
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,165
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,164
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/085,925
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/094,657
PRIOR FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1245
SOFTWARE: Patentn Ver. 2.0
SEQ ID NO 422
LENGTH: 704
TYPE: PRT
ORGANISM: Homo sapiens
US-09-933-767-422

Query Match      42.2%; Score 49; DB 12; Length 704;
```


Best Local Similarity 52.9%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSTLEAWSNDPOYNKD 21
Db 191 IMVLGASFPDPOYNKD 207

RESULT 14
US-10-023-282-422
; Sequence 422, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 422
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-282-422

Query Match 42.2%; Score 49; DB 15; Length 704;
Best Local Similarity 52.9%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSTLEAWSNDPOYNKD 21
Db 191 IMVLGASFPDPOYNKD 207

RESULT 15
US-10-156-761-9158
; Sequence 9158, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMDRA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HAYTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9158
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9158

Query Match 41.4%; Score 48; DB 15; Length 216;
Best Local Similarity 64.3%; Pred. No. 16;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HLSTLEAWSNDPQ 17

Db 184 HLSMTEAIRDRDQ 197

Search completed: November 13, 2003, 10:29:02
Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      November 13, 2003, 09:31:40 ; Search time 38.5166 Seconds
              (without alignments)
              86.541 Million cell updates/sec
```

```
Title: US-09-290-049A-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDNDTPYLHD 21
```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

Database :

1. A_Geneseq_19Jun03:*

2. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1960.DAT:*

3. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*

4. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*

5. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*

6. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*

7. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*

8. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*

9. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*

10. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*

11. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*

12. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*

13. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*

14. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*

15. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*

16. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*

17. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*

18. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*

19. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*

20. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*

21. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*

22. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*

23. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*

24. /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	119	100.0	1592	14	AAK32295	Glucoamylase
2	108	90.8	1017	23	AAU79285	Strepococcus mutans
3	108	90.8	1475	23	AAU98007	S. mutans glucosyl
4	108	90.8	1475	23	AAU98033	S. mutans glucosyl
5	108	90.8	1475	23	AAU98031	S. mutans glucosyl
6	108	90.8	1475	23	AAU98032	S. mutans glucosyl
7	108	90.8	1475	23	AAU98033	S. mutans glucosyl
8	108	90.8	1475	23	AAU98033	S. mutans glucosyl
9	108	90.8	1475	23	AAU98035	S. mutans glucosyl

1	108	90.8	1475	23	AAU98036	S. mutans glucosyl
2	108	90.8	1475	23	AAU98037	S. mutans glucosyl
3	108	90.8	1475	23	AAU98038	S. mutans GRP mut
4	108	90.8	1475	23	AAU98039	S. mutans glucosyl
5	108	90.8	1475	23	AAU98040	S. mutans glucosyl
6	108	90.8	1476	23	AAU98028	S. mutans glucosyl
7	99	83.2	1375	23	AAU98028	Streptococcus mut
8	99	83.2	1375	23	AAU98028	Streptococcus mut
9	99	83.2	1375	23	AAU98029	S. mutans glucosyl
10	99	83.2	1375	23	AAU98029	S. mutans glucosyl
11	74	62.2	1430	23	AAU98041	S. mutans glucosyl
12	74	62.2	1430	23	AAU98042	S. mutans glucosyl
13	74	62.2	1430	23	AAU98043	S. mutans glucosyl
14	74	62.2	1430	23	AAU98044	S. mutans glucosyl
15	74	62.2	1430	23	AAU98045	S. mutans glucosyl
16	72	60.5	2835	23	ABB98574	Dextran saccharase
17	72	60.5	2835	23	ABB98574	Dextran saccharase
18	66	55.5	12	23	ABB98642	Dextran saccharase
19	66	55.5	12	23	ABB98642	Leuconostoc mesent
20	54.6	54.6	1527	23	AAU80055	Dextran-saccharas
21	50.4	50.4	12	23	ABB98641	Alpha-D-glucosyltr
22	57	47.9	1577	17	AAU91047	L. mesenteroides a
23	57	47.9	1577	17	AAU10667	Dextran-saccharas
24	53	44.5	12	23	ABB98663	Arabidopsis thalia
25	53	44.5	195	21	AAU09941	Dextran saccharase
26	53	44.5	195	21	ABB98573	T. thermophilus nt
27	52	43.7	885	20	AAU75420	Drosophila melanog
28	51	42.9	1195	20	ABB59363	Human polypeptide
29	50	42.0	1252	22	AAU43544	Lactobacillus reut
30	48.5	40.8	401	22	AAU74519	Dextrane saccharas
31	48	40.3	1781	23	ABB98578	Dextrane-saccharas
32	47	39.5	12	23	ABB98644	Dextrane-saccharas
33	47	39.5	12	23	ABB98646	Dextrane-saccharas
34	47	39.5	12	23	ABB98648	LDL receptor bindi
35	47	39.5	739	22	AAU04873	Novel central nerv
36	46.5	39.1	549	22	AAU87244	Amino acid sequenc
37	46.5	39.1	1604	23	AAU82715	Lamona intracellul
38	46	38.7	486	22	AAU86552	Drosophila melanog
39	46	38.7	880	22	ABB61809	

ALIGNMENTS

RESULT 1
AA332925
AA332925 standard; Protein; 1592 AA.
AC
AA332925;
XX
DT
28-JUN-1993 (first entry)
XX
DE
Glucosyltransferase I.
XX
KW
GT-1; Streptococcus; dental; caries.
XX
OS
Streptococcus sobrinus.
XX
PN
JP05023188-A.
XX
PD
02-FEB-1993.
XX
PF
25-JUL-1991; 91JP-0186592.
XX
PR
25-JUL-1991; 91JP-0186592.
XX
PA
(FUKU/) FUKUI I.
PA
(KATO/) KATO K.
XX
DR
WPI; 1993-079449/10.
XX
DR
N-PSDB; AAQ37760.
XX
PT
DNA sequence glucosyl:transferase-I - comprises Streptococcus
PT
sobrinus DNA sequence with at least one nucleotide added or
deleted
XX

PS Claim 13; Page 15; 29pp; Japanese.

XX The DNA sequence from Streptococcus sobrinus strain 6715 encodes
 CC glucosyltransferase-I (and mutants). The DNA was obtd. by treating
 CC S. sobrinus 6715 with mutanolysin, extracting the chromosomal DNA,
 CC partially digesting with Sau3AI and fractionating on agarose gel.
 CC The 3-5 kbp fragment was ligated into pUC18 and B. coli JM109
 CC transformed with it. A GT-1 expressing clone was isolated and
 CC sequenced. The clone may be used in the development of a drug for
 CC dental caries.

XX Sequence 1592 AA;

XX Query Match 100.0%; Score 119; DB 14; Length 1592;
 XX Best Local Similarity 100.0%; Pred. No. 6.4e-09;
 XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 DB 477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 2
 AAU79285
 ID AAU79285 standard; Protein, 1017 AA.
 XX
 AC AAU79285;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 XX Streptococcus mutans monoclonal antibody-related protein #2.
 DE Streptococcus mutans monoclonal antibody-related protein #2.
 XX Antibody; dental caries; water insoluble glucan synthetase;
 KM anti-carries; glucosyl transferase-B; immunotherapy.
 XX
 XX Streptococcus mutans.
 OS JP2002114709-A.
 XX
 PN 16-APR-2002.
 XX
 PD 04-OCT-2000; 2000JP-0304899.
 XX
 PF 04-OCT-2000; 2000JP-0304899.
 XX
 PR 04-OCT-2000; 2000JP-0304899.
 XX
 PA (UNIV-) UNIV NIPPON.
 XX
 XX WPI; 2002-448101/48.
 DR
 XX
 PT Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 XX Claim 4; Page 17-19; 28pp; Japanese.
 PS
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 CC Sequence 1017 AA;
 SQ
 XX Query Match 90.8%; Score 108; DB 23; Length 1017;
 XX Best Local Similarity 85.7%; Pred. No. 1.7e-07;
 XX Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21

DB 447 ANNHVSIVEAMSDNDTPYLHD 467

RESULT 3
 AAU98027
 ID AAU98027 standard; Protein, 1475 AA.
 XX
 AC AAU98027;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB.
 DE
 XX glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture.
 XX
 XX Streptococcus mutans.
 OS US2002031826-A1.
 XX
 PN 14-MAR-2002.
 XX
 PD 19-DEC-2000; 2000US-0740274.
 XX
 PF 11-DEC-1998; 98US-0210361.
 XX
 PR 07-JUN-1995; 95US-0478704.
 XX
 PR 07-JUN-1995; 95US-0482711.
 XX
 PR 07-JUN-1995; 95US-0485243.
 XX
 PR 16-JAN-1998; 98US-0007999.
 XX
 PR 16-JAN-1998; 98US-0008172.
 XX
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 PI
 XX WPI; 2002-414332/44.
 DR
 DR N-PSDB; ABR52938.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Disclosure; Page 21-25; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571X, D567T/D571X,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents GTFB.
XX
SQ Sequence 1475 AA;
Query Match 90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAWSDNDPYLHD 21
DB 481 ANDHSLTEAWSDNDPYLHD 501
RESULT 4
AAU98030 standard; Protein; 1475 AA.
XX AAU98030;
XX 27-AUG-2002 (first entry)
XX S. mutans glucosyltransferase GTFB mutant I448V.
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutein.
XX Streptococcus mutans.
XX Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 448 /note= "Wild-type Ile substituted by Val"
XX FT
XX US2002031826-A1.
XX PD 14-MAR-2002.
XX 19-DEC-2000; 2000US-0740274.
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX (NICH// NICHOLS S E.
XX PA
XX PI Nichols SE;
XX WPI, 2002-414332/44.
XX DR
XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substitutes for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
XX (GTF) B polypeptide having changes at position from I448V, D457N,
XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC

CC Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N477D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper string and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAWSDNDPYLHD 21
DB 481 ANDHSLTEAWSDNDPYLHD 501

RESULT 5
AAU98031 standard; Protein; 1475 AA.
XX AAU98031;
XX 27-AUG-2002 (first entry)
XX S. mutans glucosyltransferase GTFB mutant D457N.
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutein.
XX Streptococcus mutans.
XX Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
XX FT
XX US2002031826-A1.
XX PD 14-MAR-2002.
XX 19-DEC-2000; 2000US-0740274.
XX

XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX MPI, 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing the glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 SQ

Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIWEASGNDPTPLAD 21
 ||:||||:||||:||||:
 Db 481 ANDHUSILWEASGNDPTPLAD 501

RESULT 6

AAU98032
 ID AAU98032 standard; Protein, 1475 AA.
 XX
 AC AAU98032;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muteln.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 OS
 FT Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX MPI, 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing the glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GTF, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GTF mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GTFB sequence appearing as AAU98027
CC	and the information in claim 36.
XX	
SQ	Sequence 1475 AA;
Query Match	90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity	85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative	3; Mismatches 0; Indels 0; Gaps 0;
QY	1 ANNHVSIVGANSNDTPYLHD 21
	: : : : : : : : :
Db	481 ANDHLSIIEAMSDNDTPYLHD 501
RESULT 7	
ID	AAU98033
AC	AAU98033 standard; Protein; 1475 AA.
XX	
AC	AAU98033;
DT	27-AUG-2002 (first entry)
XX	
DE	S. mutans glucosyltransferase GTFB mutant K1014T.
XX	
KM	Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amyloplast; vacuole; paper manufacture; mutant; mutcin.
XX	
OS	Streptococcus mutans.
OS	Synthetic.
XX	
EH	Key Location/Qualifiers
FT	Misc-difference 1014
FT	/note= "Wild-type Lys substituted by Thr"
XX	
PN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 98US-0007959.
PR	16-JAN-1998; 98US-0008172.
PR	20-JAN-1998; 98US-0009620.
XX	
PA	(NICH/) NICHOLS S E.
XX	
PI	Nichols SE;
XX	
DR	WPI; 2002-414332/44.
XX	
PT	Glucosyltransferase B or D protein useful for producing a glucan useful
PT	as substitutes for and additions to modified starch and latexes in
PT	paper manufacture, comprises mutations in specific positions -
XX	
PS	Claim 36; Page -; 44pp; English.
XX	
CC	The invention an isolated protein comprising a glucosyltransferase
CC	(GTF) B polypeptide having changes at position from 1448V, D457N,

[illegible]

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 XX WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 PS The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, paper
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX Sequence 1475 AA;
 SQ Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pxed. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVFAMSDNDPFLYLD 21
 ||:|||||

DB 481 ANNHVSIVFAMSDNDPFLYLD 501
 RESULT 9
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.
 XX AC AAU98035;
 XX AC
 DT 27-AUG-2002 (first entry)
 XX DE S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 XX DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX OS
 OS Streptococcus mutans.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 XX US2002031826-A1.
 PN 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 XX WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 PS The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the expression cassette, host cell
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

SO Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAMSDNDPYLHD 21
 DB 481 ANDHSTLEAMSDNDPYLHD 501

RESULT 10

AAU98036 standard; Protein; 1475 AA.

AC AAU98036;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KW amyloplast; vacuole; paper manufacture; mutant; mutain.

OS Streptococcus mutans.

OS Synthetic.

FT Key Location/Qualifiers

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FN US2002031826-A1.

PN 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.

PE 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX MPI; 2002-414332/44.

PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y163N/Y170K/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

SO Sequence 1475 AA;

Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAMSDNDPYLHD 21
 DB 481 ANDHSTLEAMSDNDPYLHD 501

RESULT 11

AAU98037 standard; Protein; 1475 AA.

AC AAU98037;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KW amyloplast; vacuole; paper manufacture; mutant; mutain.

OS Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI MPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169N/Y170N/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex, where
 CC the thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize like
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU96027

CC and the information in claim 36.
 XX SQ Sequence 1475 AA;
 XX Query Match 90.8%; Score 108; DB 23; Length 1475;
 XX Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 XX Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 ANNHVSIVKAWSDNDPRLAD 21
 XX ||:||||:||||:||||:||||:
 XX DB 481 ANDHSLILEKWSNDPRLAD 501
 XX
 XX RESULT 12
 XX ID AAU98038 standard; Protein; 1475 AA.
 XX AC AAU98038;
 XX DT 27-AUG-2002 (first entry)
 XX DE S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
 XX KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX KW coating composition; glucan; starch; latex; thermoplastic molecule;
 XX KW amyloplast; vacuole; paper manufacture; mutant; mutcin.
 XX OS Streptococcus mutans.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Ile substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI MPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, I48V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;
 XX

Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAMSDNDTPYLHD 21
 ||:|||||
 DB 481 ANDHLSIEAMSDNDTPYLHD 501

RESULT 13
 ID AAU98039 standard; Protein; 1475 AA.
 XX
 AC AAU98039;
 XX

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.
 XX

OS Streptococcus mutans.
 XX Synthetic.
 XX

XX Key Location/Qualifiers
 FH Misc-difference 169..171

FT /note= "Wild-type Tyr-Tyr-Tyr substituted by
 FT Ala-Ala-Ala"
 XX
 XX

PN US2002031826-A1.

XX
 PD 14-MAR-2002.

XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 XX
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JUN-1998; 98US-0007999.
 PR 16-JUN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-41432/44.

PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -: 44p; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I48V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I48V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;
 XX

Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAMSDNDTPYLHD 21
 ||:|||||
 DB 481 ANDHLSIEAMSDNDTPYLHD 501

RESULT 14
 AAU98040
 ID AAU98040 standard; Protein; 1475 AA.
 XX
 AC AAU98040;
 XX
 DT 27-AUG-2002 (first entry)
 DE
 XX S. mutans glucosyltransferase GTFB mutant K7790.
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 779 /note= "wild-type Lys substituted by Gln"
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K7790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 XX
 QY Query March 90.8%; Score 108; DB 23; Length 1475;
 DB Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAWSNDPTPLHD 21
 DB 481 ANNHVSIVEAWSNDPTPLHD 501
 XX
 RESULT 15
 AAU9284
 ID AAU9284 standard; Protein; 1476 AA.
 XX
 AC AAU9284;
 XX
 DT 13-AUG-2002 (first entry)
 DE
 XX Streptococcus mutans monoclonal antibody-related protein #1.
 XX
 KM Antibody; dental caries; water insoluble glucan synthetase;
 KM anti-caries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 PN JP2002114709-A.
 XX
 PD 16-APR-2002.
 XX
 PF 04-OCT-2000; 2000JP-0304889.
 XX
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PA (UYN-) UNITV NIPPON.
 XX
 DR WPI; 2002-448101/48.
 XX
 PT Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 PS Claim 3; Page 13-16; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX
 SQ Sequence 1476 AA;
 XX
 QY Query March 90.8%; Score 108; DB 23; Length 1476;

Thu Nov 13 12:12:07 2003

us-09-290-049a-12.rag

Page 11

Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0.

```
QY      1 ANNHVSIVEAWSNDNTPYLLHD 21
      ||::||::||::||::||::||
Db      481 ANDHLSILEAWSNDNTPYLLHD 501
```

Search completed: November 13, 2003, 09:38:27
Job time : 39.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDPTPLYLD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	119	100.0	A38175	glucosyltransferase
2	108	90.8	A38175	glcB protein precu
3	99	83.2	JT0345	dextranucrase (EC
4	74	62.2	A45866	dextranucrase (EC
5	68	57.1	T31098	probable dextranu
6	67	56.3	A41483	glucosyltransferas
7	61	51.3	T30857	glucosyltransferas
8	61	51.3	T30552	glucosyltransferas
9	57	47.3	T30858	glucosyltransferas
10	56	47.1	A44811	glucosyltransferas
11	55	46.2	S22737	hypothetical prote
12	53	44.5	C86205	BglF2 protein - nu
13	49	41.2	Q08E40	conserved hypothet
14	49	41.2	B82220	glycerinaldehyde-3-p
15	48	40.3	A38175	hypothetical prote
16	46	38.7	S59448	hypothetical prote
17	46	38.7	H70538	probable pdk prot
18	46	38.7	H70538	pyruvate, phosphat
19	45.5	38.2	D90506	4-aminobutyrate am
20	45	37.8	C64233	hypothetical prote
21	45	37.8	T48610	Rogger domain cont
22	45	37.8	T40088	hypothetical prote
23	45	37.8	E86362	hypothetical prote
24	45	37.8	T02748	pristinamycin I sy
25	45	37.8	T30289	site-specific DNA-
26	44.5	37.4	835	metalloproteinase
27	44	37.0	A59414	X-Pro dipeptidyl-p
28	44	37.0	UC5142	probable membrane
29	44	37.0	S64443	

30	44	37.0	1131	2	T16217	hypothetical prote
31	44	37.0	1313	2	T29027	hypothetical prote
32	43.5	36.6	491	1	D64947	glucose-6-phosphat
33	43.5	36.6	491	2	B90349	glucose-6-phosphat
34	43.5	36.6	491	2	F85797	glucose-6-phosphat
35	43.5	36.6	491	2	AB0742	hypothetical prote
36	43	36.1	78	2	S76593	conserved hypothet
37	43	36.1	194	2	H72037	CT647 hypothetical
38	43	36.1	194	2	C66586	hypothetical prote
39	43	36.1	275	2	S55978	heat-shock sigma f
40	43	36.1	284	2	JC7148	glycerinaldehyde-3-p
41	43	36.1	331	1	DEUTGC	hypothetical prote
42	43	36.1	346	2	E70715	hypothetical prote
43	43	36.1	351	2	D89991	probable glutathio
44	43	36.1	396	2	T35024	probable homogenit
45	43	36.1	441	2	T36737	

ALIGNMENTS

RESULT 1
A38175
glucosyltransferase precursor - Streptococcus sobrinus
C/Species: Streptococcus sobrinus
C/Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
R/Accession: A38175
R/Abn, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
U. Bacteriol. 173, 989-996, 1991
A>Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A/Reference number: A38175; MUID:91123227; PMID:1704006
A/Accession: A38175
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1592 <ABO>
C/Superfamily: cpl repeat homology
A/Cross-references: GB:D90213; NID:9217032; PIDN:BA14241.1; PID:dl014946; PID:9217033
F/1093-1112/Domain: cpl repeat homology <CP1>
F/1222-1241/Domain: cpl repeat homology <CP2>
F/1287-1306/Domain: cpl repeat homology <CP3>
F/1330-1351/Domain: cpl repeat homology <CP4>
F/1352-1371/Domain: cpl repeat homology <CP5>
F/1402-1420/Domain: cpl repeat homology <CP6>
F/1465-1484/Domain: cpl repeat homology <CP7>
F/1513-1532/Domain: cpl repeat homology <CP8>

Query Match 100.0%; Score 119; DB 2; Length 1592;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLYLD 21
DB 477 ANNHVSIVEAWSNDPTPLYLD 497

RESULT 2
B33135
glcB protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
R/Accession: B33135
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
U. Bacteriol. 169, 4263-4270, 1987
A>Title: Sequence analysis of the glcB gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Accession: B33135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1475 <SHI>
A/Cross-references: GB:M17361; NID:g153639; PIDN:AAA86588.1; PID:g153640
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990
A/Reference number: A33128

A:Accession: A33128
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-171,173-661,'N',643-1475 <SH2>
 A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 90.8%; Score 108; DB 2; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.6e-08;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 3
 J:Gen. Microbiol. 136, 2099-2105, 1990
 A:Title: Nucleotide sequence of the Streptococcus mutans glfD gene encoding the glucosyl
 A:Reference number: A45866; PMID:2148600
 A:Accession: A45866
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1431 <HON>
 A:Cross-references: GB:M29296
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase, hexosyltransferase
 F:181-201/Domain: cpl repeat homology <CP1>
 F:1127-1146/Domain: cpl repeat homology <CP2>
 F:1192-1211/Domain: cpl repeat homology <CP3>
 F:1257-1276/Domain: cpl repeat homology <CP4>
 F:1277-1297/Domain: cpl repeat homology <CP5>
 F:1321-1340/Domain: cpl repeat homology <CP6>
 F:1341-1361/Domain: cpl repeat homology <CP7>
 F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 62.2%; Score 74; DB 2; Length 1431;
 Best Local Similarity 66.7%; Pred. No. 0.005;
 Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 495 ANNHLSILEAMSDNDTPYLHD 515

RESULT 5
 T31098
 probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T31098
 R:Monchois, V.; Remaud-Simeon, M.; Monsar, P.; Willemot, R.M.
 FEMS Microbiol. Lett. 159, 307-315, 1998
 A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A:Reference number: Z20981; PMID:98164374; PMID:9503626
 A:Accession: T31098
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1508 <MON>
 A:Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AA95453.1
 A:Experimental source: strain NRRL B-1299
 C:Genetics:
 A:Gene: dsrB
 C:Function:
 A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C:Keywords: glycosyltransferase, hexosyltransferase

Query Match 57.1%; Score 68; DB 2; Length 1508;
 Best Local Similarity 57.1%; Pred. No. 0.046;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 563 ANQHLSILEAMSHNDPEYKVD 583

RESULT 6
 A41483
 glucosyltransferase (EC 2.4.1.-) glfs precursor - Streptococcus sobrinus
 C:Species: Streptococcus sobrinus
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C:Accession: A41483
 R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
 Infect. Immun. 58, 2452-2458, 1990
 A:Title: Analysis of the Streptococcus downei glfs gene, which specifies a glucosyltran-
 A:Reference number: A41483; PMID:90316655; PMID:2142479
 A:Accession: A41483
 A:Molecule type: DNA
 A:Residues: 1-1365 <GLI>
 A:Cross-references: GB:M30943; NID:g153652; PIDN:AAA6898.1; PID:g153653

C:Genetics:
A:Gene: gtfS
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 56.3%; Score 67; DB 2; Length 1365;
Best Local Similarity 57.1%; Pred. No. 0.058;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVFAMSDNDTPYLHD 21
DB 467 AIDHLSTLEAMSGNDNDYKDD 487

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30857

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <SIM>
A:Cross-references: EMBL:LJ5495; NID:G662378; PID:G662379; PIDN:AAC41412.1
A:Gene: gtfL

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVFAMSDNDTPYLHD 21
DB 536 AIKHLSTLEAMSGNDNDYKDD 556

RESULT 8
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)

C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaffe, R.I.
submitted to the EMBL Data Library, February 1998

A:Description: Streptococcus salivarius V1477 gtfN.
A:Reference number: Z20854
A:Accession: T30552

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <JAF>
A:Cross-references: EMBL:AF049609; NID:G2935545; PID:G2935546; PIDN:AAC05156.1
A:Gene: gtfN

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVFAMSDNDTPYLHD 21
DB 536 AIKHLSTLEAMSGNDNDYKDD 556

RESULT 9
T30858
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>

A:Cross-references: EMBL:LJ5928; NID:G662380; PID:G662381; PIDN:AAC41413.1
A:Gene: gtfM

Query Match 47.9%; Score 57; DB 2; Length 1577;
Best Local Similarity 61.1%; Pred. No. 2.5;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 HVSIVFAMSDNDTPYLHD 21
DB 591 HSLTLEAMSGNDNDYKDD 608

RESULT 10
A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius

C:Species: Streptococcus salivarius
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811

A:Molecule type: DNA
A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:Z11873; NID:G47526; PIDN:CAAT7900.1; PID:G47527

A:Note: Sequence extracted from NCBI backbone (NCBI:G1050, NCBI:P:G1052)
A:Gene: gtfL
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 47.1%; Score 56; DB 2; Length 1518;
Best Local Similarity 60.0%; Pred. No. 3.4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 HVSIVFAMSDNDTPYLHD 18
DB 504 HVSIVFAMSLNDNHV 518

RESULT 11
S22727
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C:Species: Streptococcus salivarius
C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C:Accession: S22727; S28810; B44811; S22727
R:Jacques, N.
submitted to the EMBL Data Library, March 1992

A:Reference number: S22726
A:Accession: S22737
A:Molecule type: DNA
A:Residues: 1-1599 <JAC>

A:Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAAT7898.1; PID:G47531
A:Experimental source: ATCC 25975
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: S28810
A:Molecule type: DNA
A:Residues: 1-51 <GIF>

A;Cross-references: EMBL:Z11873

A;Genetics:

A;Gene: gtlX

C;Superfamily: cpl repeat homology

C;Keywords: glycosyltransferase; hexosyltransferase

F;1456-1475/Domain: cpl repeat homology <CPR>

Query Match 46.2%; Score 55; DB 2; Length 1599;

Best Local Similarity 50.0%; Pred. No. 5.2;

Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 HVSIVAEWSDNDTPYLHD 21

Db 494 NISLIEAWSHNDPIYVNE 511

RESULT 12

C66205 hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: C66205

R;Theologos, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.E.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberrg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C66205

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1175 <STO>

A;Cross-references: GB:AE005172; NID:g9954041; PIDN:AAF62215.1; GSPDB:GN00.41

C;Genetics:

A;Map position: 1

Query Match 44.5%; Score 53; DB 2; Length 175;

Best Local Similarity 42.1%; Pred. No. 0.78;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHVSIVAEWSDNDTPYLHD 21

Db 109 NHOEVIDAMSDHOKPLMTD 127

RESULT 13

BGLR2 protein - human herpesvirus 4 (strain B95-8)

C;Species: human herpesvirus 4, Epstein-Barr virus

C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999

C;Accession: C43044; J01381; A03784; A03794; S33036

R;Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrall, B.G.

Mol. Biol. Med. 1, 21-45, 1983

A;Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus

A;Reference number: A93065; MUID:85035713; PMID:6092825

A;Accession: C43044

A;Molecule type: DNA

A;Residues: 1-336 <BAN>

A;Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24831.1; PID:g1334895

R;Beier, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H

Nature 310, 207-211, 1984

A;Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A;Reference number: A03794; MUID:84270667; PMID:6087149

A;Contents: annotation; protein coding region

A;Note: neither amino acid nor nucleotide sequence is given

R;Chen, M.R.; Hsu, T.Y.; Lin, S.W.; Chen, J.Y.; Yang, C.S.

J. Gen. Virol. 72, 3047-3055, 1991

A;Title: Cloning and characterization of cDNA clones corresponding to transcripts from B

A;Reference number: J01381; MUID:92113548; PMID:1662696

A;Accession: J01381

A;Molecule type: mRNA

A;Residues: 1-336 <CHE>

A;Cross-references: GB:S77132; NID:g243314; PIDN:AAB21113.1; PID:g243315

C;Superfamily: Epstein-Barr virus BGLF2 protein

Query Match 41.2%; Score 49; DB 1; Length 336;

Best Local Similarity 41.2%; Pred. No. 7.1;

Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 ANNHVSIVAEWSDNDTP 17

Db 216 AGAHVNILRGMTEDDSP 232

RESULT 14

D82220 conserved hypothetical protein VCI268 [imported] - Vibrio cholerae (strain N16961 serogr

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: D82220

R;Heidelberg, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Charlson, D.; Ermolaeva, M.D.; Vamathevan, T.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82220

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-524 <HEI>

A;Cross-references: GB:AE004206; GB:AE003852; NID:g9655749; PIDN:AAF94427.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCI268

A;Map position: 1

Query Match 41.2%; Score 49; DB 2; Length 524;

Best Local Similarity 43.5%; Pred. No. 12;

Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 ANNHVSIVAEWSDNDTP 21

Db 365 ANNHWEILRGMSDPEVIDPALD 387

RESULT 15

B48445 glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me

C;Species: Leishmania mexicana

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002

C;Accession: B48445; S25142

R;Hanaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Opperdoes, F.R.; Michels, P.A.M.

Mol. Biochem. Parasitol. 55, 115-126, 1992

A;Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate d

A;Reference number: A48445; MUID:93063042; PMID:1435864

A;Accession: B48445

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-331 <HAN>

A;Cross-references: EMBL:X65220; NID:g9552; PIDN:CAA46323.1; PID:g9553

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C;Keywords: oxidoreductase

Query Match 40.3%; Score 48; DB 2; Length 331;

Best Local Similarity 42.1%; Pred. No. 9.9;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 NNHVSIVAEWSDNDTPYLH 20

Db 301 NDHFVKLVSWYDNETGSH 319

Thu Nov 13 12:12:08 2003

us-09-290-049a-12.rpr

Page 5

Search completed: November 13, 2003, 09:50:26
Job time : 14.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-12

Perfect score: 119
Sequence: 1 ANNHVSIYKMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	119	100.0	1 GTF2_STRDO	P27470 streptococ
2	119	100.0	1 GTF1_STRDO	P11001 streptococ
3	108	90.8	1 GTFB_STRMU	P08987 streptococ
4	99	83.2	1 GTRC_STRMU	P13470 streptococ
5	74	62.2	1 GTFD_STRMU	P49331 streptococ
6	67	56.3	1 GTRF_STRDO	P29336 streptococ
7	49	41.2	1 GTRF_STRDO	P03221 epstein-bar
8	48	40.3	1 G3PC_LEIME	O01558 leishmania
9	46	38.7	1 YMS8_YEAST	O03695 saccharomyc
10	45	37.8	1 G3P_MYCGR	P47543 mycoplasma
11	44	37.0	1 YG3C_YEAST	O92093 mus musculu
12	44	37.0	1 YG3C_YEAST	P52280 saccharomyc
13	44	37.0	1 YPO4_CAEEL	P53015 caenorhadi
14	44	37.0	1 PCGV_MOUSE	O62059 mus musculu
15	43.5	36.6	1 G6PD_ECOLI	P22992 escherichia
16	43	36.1	1 SC35_YEAST	P53271 saccharomyc
17	43	36.1	1 G3PC_TRYRB	P10097 trypanosoma
18	43	36.1	1 Y943_MYCTU	P71566 mycobacteri
19	43	36.1	1 HBD_STRCO	O92285 streptomyc
20	43	36.1	1 DOPO_HUMAN	P09172 homo sapien
21	43	36.1	1 SUHM_DROYI	O08876 drosophila
22	42.5	35.7	1 AMP2_STRCO	O60394 streptomyc
23	42.5	35.7	1 G6PD_STRCH	P37986 erwinia chr
24	42	35.3	1 P6F1_DROME	O92637 drosophila
25	42	35.3	1 CG16_YEAST	P20867 saccharomyc
26	42	35.3	1 G3P_BACST	P00362 bacillus st
27	42	35.3	1 G3P_BACCO	P11115 bacillus co
28	42	35.3	1 HGD_CAUCR	O94518 caulobacter
29	42	35.3	1 HGD_PSEAR	O94590 pseudomonas
30	42	35.3	1 GARD_PYRUF	O80400 pyrococcus
31	42	35.3	1 SYM_UREPA	O92906 ureaplasma
32	42	35.3	1 YE09_MYCPN	P73575 mycoplasma
33	42	35.3	1 AD28_HUMAN	O92906 ureaplasma

34	42	35.3	776	1	AD28_MYCPN	O92906 ureaplasma
35	42	35.3	1627	1	ADP1_MYCPN	P11311 mycoplasma
36	42	35.3	3898	1	POIG_HCVB	P21530 hog cholera
37	41.5	34.9	416	1	HGD_LEGPN	O94400 legionella
38	41.5	34.9	468	1	KG3H_DICDI	P51136 dictyostell
39	41.5	34.9	578	1	VAC8_YEAST	P39568 saccharomyc
40	41.5	34.9	684	1	ISH1_SCHPO	O92746 schizosacch
41	41	34.5	130	1	UCR6_SCHMA	O01374 schistosoma
42	41	34.5	236	1	UCR4_TOBAC	P51134 nicotiana t
43	41	34.5	241	1	HEP1_HAEIN	P35757 haemophilus
44	41	34.5	241	1	HEP2_HAEIN	P45991 haemophilus
45	41	34.5	301	1	Y186_MYCPN	P75265 mycoplasma

ALIGNMENTS

```

RESULT 1
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; Pubmed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RT J. Bacteriol. 173:989-996(1991).
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N-1) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N-1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D90213; BAA14241.1;
CC InterPro: IPR002479; CM binding.
CC InterPro: IPR003118; Glyco_hydro_70.
CC Pfam: PF01473; CM_binding_1.13.
CC Pfam: PF03324; Glyco_hydro_70; 1.
CC Trasnferase; Glucosyltransferase; Signal; Repeat; Dental caries.
CC SIGNAL 1 38 POTENTIAL.
CC CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
CC FT 39 1592 CATALYTIC (APPROXIMATE).
CC FT DOMAIN 39 1044

```

FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
 FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
 FT REPEAT 1093 1142 1.
 FT REPEAT 1158 1207 2.
 FT REPEAT 1222 1272 3.
 FT REPEAT 1287 1337 4.
 FT REPEAT 1402 1451 5.
 FT REPEAT 1514 1563 6.
 FT REPEAT 1577 1592 7 (INCOMPLETE).
 SQ SEQUENCE 1592 AA; 176167 MW; BC0A6D079351ECF CRC64;
 Query Match 100.0%; Score 119; DB 1; Length 1592;
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 2
 GTF1_STRDO STANDARD; PRT; 1597 AA.
 ID GTF1_STRDO
 AC P11001;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFI.
 OS Streptococcus downei (Streptococcus sobrius).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MFE28;
 RC MEDLINE=87308014; PubMed=3040686;
 RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
 RA "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
 RT sobrius MFE28."
 RL J. Bacteriol. 169:4271-4278(1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
 CC fructose + (1,6)-alpha-D-glucosyl (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S.MUTANS.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>;
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL, M17391; AAC63063.1; -;
 CC InterPro; IPR002479; CW binding.
 CC InterPro; IPR003131; Glyco_hydro_70.
 CC Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 38
 FT POTENTIAL.

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
 FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
 FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
 FT REPEAT 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
 FT REPEAT 1099 1132 A REPEAT.
 FT REPEAT 1163 1213 AC REPEAT.
 FT REPEAT 1227 1277 AC REPEAT.
 FT REPEAT 1292 1342 AC REPEAT.
 FT REPEAT 1352 1399 B REPEAT.
 FT REPEAT 1406 1455 AC REPEAT.
 FT REPEAT 1465 1512 B REPEAT.
 FT REPEAT 1519 1568 AC REPEAT.
 FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
 SQ SEQUENCE 1597 AA; 177080 MW; B9E6A200868798B CRC64;
 Query Match 100.0%; Score 119; DB 1; Length 1597;
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 483 ANNHVSIVEAMSDNDTPYLHD 503

RESULT 3
 GTFB_STRMU STANDARD; PRT; 1476 AA.
 ID GTFB_STRMU
 AC P08987; O69381; O69384; O69387; O69390; O69396;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFB OR SMU.1004.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RC MEDLINE=87308013; PubMed=3040685;
 RA Shirota T., Ueda S., Kuramitsu H.K.;
 RA "Sequence analysis of the gtfb gene from Streptococcus mutans."
 RL J. Bacteriol. 169:4263-4270(1987).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=MT429 / Serotype C; MT4245 / Serotype E, MT4251 / Serotype F,
 CC MT4467 / Serotype B, and MT8148 / Serotype C;
 CC MEDLINE=98231643; PubMed=9570124;
 CC RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
 CC Kimura S., Hamada S.;
 CC "Molecular analyses of glucosyltransferase genes among strains of
 CC Streptococcus mutans."
 CC FEMS Microbiol. Lett. 161:331-336(1998).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN=UA159 / ATCC 700610 / Serotype C;
 CC MEDLINE=22295063; PubMed=12397186;
 CC RA Ajdic D., Moshan M.M., McLaughlin R.E., Savic G., Chang J.,
 CC Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 CC Li S., Zhu H., Nislar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 CC "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 CC pathogen."
 CC Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
 CC fructose + (1,6)-alpha-D-glucosyl (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.

```

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCAN.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17361; AAA8588.1; -
DR EMBL; D88651; BAA26101.1; -
DR EMBL; D88654; BAA26105.1; -
DR EMBL; D88657; BAA26109.1; -
DR EMBL; D88660; BAA26113.1; -
DR EMBL; D89977; BAA26119.1; -
DR EMBL; AE014940; AAN58705.1; -
DR InterPro; IPR002479; CM binding.
DR InterPro; IPR003318; Glyco hydro_70.
DR Pfam; PF01473; CM-binding_1.1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
KM Complete proteome.
PT SIGNAL 1 34
PT CHAIN 35 1476
PT DOMAIN 35 1051
PT DOMAIN 1097 1476
PT REPEAT 1097 1130
PT DOMAIN 1161 1470
PT REPEAT 1161 1210
PT REPEAT 1225 1275
PT REPEAT 1290 1340
PT REPEAT 1355 1405
PT REPEAT 1420 1470
PT REPEAT 62 62
PT REPEAT 65 65
PT REPEAT 68 68
PT VARIANT 78 78
PT VARIANT 86 86
PT VARIANT 89 89
PT VARIANT 168 168
PT VARIANT 276 276
PT VARIANT 399 399
PT VARIANT 474 474
PT VARIANT 512 512
PT VARIANT 519 519
PT VARIANT 701 701
PT VARIANT 708 708
PT VARIANT 938 938
PT VARIANT 952 952
PT VARIANT 963 964
PT VARIANT 968 970
PT VARIANT 1086 1086
PT VARIANT 1158 1158
PT VARIANT 1163 1163
PT VARIANT 1168 1168
PT VARIANT 1182 1182
PT VARIANT 1234 1234
PT VARIANT 1263 1263
PT VARIANT 1264 1264
PT VARIANT 1264 1264

S -> T (IN STRAIN MT4239).
T -> I (IN STRAIN GS-5).
V -> A (IN STRAINS GS-5, MT4245, MT4251,
MT4467 AND MT8148).
I -> S (IN STRAIN MT4251).
Q -> P (IN STRAIN MT4251).
S -> F (IN STRAIN MT4251).
K -> N (IN STRAIN MT4251).
S -> D (IN STRAINS GS-5, MT4467 AND
MT8148).
N -> R (IN STRAIN MT4239).
I -> T (IN STRAIN MT4239).
K -> R (IN STRAIN MT8148).
F -> Y (IN STRAIN MT8148).
T -> I (IN STRAIN MT8148).
A -> V (IN STRAIN MT8148).
F -> L (IN STRAIN MT8148).
FGKPYE -> YGTPVA (IN STRAINS GS-5, MT4239
AND MT4467).
SV -> NT (IN STRAINS GS-5, MT4239 AND
MT4467).
ADS -> VDG (IN STRAINS GS-5, MT4239 AND
MT4467).
A -> T (IN STRAIN MT4239).
S -> N (IN STRAIN MT4239).
H -> Y (IN STRAIN MT4251).
E -> K (IN STRAIN MT8148).
Y -> C (IN STRAIN MT8148).
A -> P (IN STRAIN MT4239).
R -> P (IN STRAIN MT8148).
R -> H (IN STRAINS GS-5 AND MT4467).
Y -> H (IN STRAINS GS-5, MT4239, MT4467

```

```

FT FT AND MT8148).
FT FT S -> G (IN STRAINS GS-5, MT4239, MT4467
FT FT AND MT8148).
FT FT H -> Y (IN STRAINS GS-5 AND MT4467).
FT FT Y -> Y (IN STRAINS GS-5, MT4239, MT4467
FT FT AND MT8148).
FT FT S -> G (IN STRAINS GS-5, MT4239, MT4467
FT FT AND MT8148).
FT FT Y -> H (IN STRAIN MT4467).
FT FT R -> A (IN REF. 1).
FT FT ADQYRVASTAPSTDGK -> LKMFALRLAPHOQMA
FT FT (IN REF. 1).
FT FT H -> L (IN REF. 1).
SQ SEQUENCE 1476 AA; 165846 MW; 906E09F731B4C8CF CRC64;

Query Match 90.8%; Score 108; DB 1; Length 1476;
Best Local Similarity 85.7%; Pred. No. 5; E-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAGSDNDPTVYHD 21
DB 481 ANDHLSILEAWSNDPTVYHD 501

RESULT 4
ID GTF-1 STRMU STANDARD; PRT; 1455 AA.
AC P13470; O69382; O69385; O69388; O69391; O69397; P05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN GTF-1 OR SMU.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gltf gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
MT4467 / Serotype B, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397185;
RA Ajdic D., McNamee W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gltf gene from Streptococcus mutans."
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT

```

Query	1 ANNHVSIVEMASDNDPTPYLHD 21	Score 99; DB 1; Length 145;
Best Local Similarity	81.0%;	Pred. No. 1.6e-07;
Matches 17; Conservative	3; Mismatches 1; Indels 0; Gaps 0	
Db	507 ANHLSILEANSYNDPTPYLHD 527	
RESULT 5		
CTFD_STRMU	STANDARD; PRT; 1462 AA.	
ID	GTFD_STRMU	
AC	P4931; 069383; 069386; 069389; 069392; 069398;	
DT	01-FEB-1996 (Rel. 33, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)	
DE	(Sucrose 6-glucosyltransferase).	
GN	GTFD OR SMU.910.	
OS	Streptococcus mutans.	
OS	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1309;	
RN	[[1]]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=GS-5;	
RC	MEDLINE=91100958; PubMed=2118600;	
RA	Honda O., Kato C., Kuramitsu H.K.;	
RT	"Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme.";	
RL	J. Gen. Microbiol. 136:2099-2105(1990).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,	
RC	MT4467 / Serotype B, and MT8148 / Serotype C;	
RA	MEDLINE=98231643; PubMed=9570124;	
RA	Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,	
RT	"Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans.";	
RL	J. Gen. Microbiol. Lett. 161:331-336(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=UA159 / ATCC 700610 / Serotype C;	
RC	MEDLINE=22295063; PubMed=12397186;	
RA	Ajdlic D., Moshan W., McLaughlin R.E., Savic G., Chang J.,	
RA	Carson M.B., Pirneman C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,	
RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perrecci J.J.;	
RT	"Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).	
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.	

CC	-1- CATABOLIC ACTIVITY. Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- DISEASE: DENTAL CARIES.
CC	-1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-S1 SYNTHESIZES BOTH FORMS OF GLUCANS.
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S.MUTANS.
CC	-1- SIMILARITY: Contains 6 cell wall binding repeats.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@ib-sib.ch).
CC	-----
DR	EMBL; M29296; AAA26895.1; -
DR	EMBL; D88653; BAA26103.1; -
DR	EMBL; D88656; BAA26107.1; -
DR	EMBL; D88659; BAA26111.1; -
DR	EMBL; D88662; BAA26115.1; -
DR	EMBL; D89979; BAA26121.1; -
DR	EMBL; AE014932; AAN58619.1; -
DR	InterPro; IPR002479; CW binding.
DR	InterPro; IPR003318; Glyco_hydro_70.
DR	Pfam; PF02324; Glyco_hydro_70; 1
KW	Transferase; Glycosyltransferase; Signal; Repeat; Dental caries; Complete proteome.
FT	FT SIGNAL 1 ?
FT	FT CHAIN 1 1462
FT	FT DOMAIN 1232 1433
FT	FT REPEAT 1232 1295
FT	FT REPEAT 1296 1359
FT	FT REPEAT 1360 1433
FT	FT REPEAT 1360 1403
FT	FT VARIANT 19 19
FT	FT VARIANT 58 58
FT	FT VARIANT 68 68
FT	FT VARIANT 81 81
FT	FT VARIANT 113 113
FT	FT VARIANT 122 122
FT	FT VARIANT 132 132
FT	FT VARIANT 135 135
FT	FT VARIANT 137 137
FT	FT VARIANT 202 202
FT	FT VARIANT 255 255
FT	FT VARIANT 275 275
FT	FT VARIANT 288 288
FT	FT VARIANT 301 301
FT	FT VARIANT 313 313
FT	FT VARIANT 317 317
FT	FT VARIANT 328 328
FT	FT VARIANT 350 350
FT	FT VARIANT 628 633
FT	FT VARIANT 688 688
FT	FT VARIANT 726 732
FT	FT VARIANT 726 730
FT	FT VARIANT 762 762
FT	FT VARIANT 964 964
FT	FT VARIANT 1019 1019

FT	VARIANT	1059	1060
FT	VARIANT	1060	1060
FT	VARIANT	1080	1080
FT	VARIANT	1142	1142
FT	VARIANT	1198	1198
FT	VARIANT	1220	1220
FT	VARIANT	1280	1280
FT	VARIANT	1282	1282
FT	VARIANT	1290	1290
FT	VARIANT	1311	1311
FT	VARIANT	1403	1403
FT	VARIANT	1425	1425
FT	VARIANT	1449	1449
FT	CONFLICT	1428	1462
SQ	SEQUENCE	1462 AA;	163387 MW; CFAA279C4D708645 CRC64;
Query Match		Best Local Similarity	62.2%; Score 74; DB 1; Length 1462;
Matches	14,	Conservative	2; Mismatches 5; Indels 0; Gaps 0;
Oy	1 ANNHVSIVEAWSNDNPYLYHD 21 : : : :		
Dn	49S AINHTSLTEAWSDNDPOYNKD 51S		
RESULT 6			
GTF5_STRDO	STANDARD;	PRT;	1365 AA.
ID_GTF5_STRDO	P29336;		
DT	01-DEC-1992 (Rel. 24, Created)		
DT	01-DEC-1992 (Rel. 24, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)		
DE	(Sucrose 6-glucosyltransferase).		
GN	GTF5.		
OS	Streptococcus downei (Streptococcus sobrinus).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
CC	Streptococcus.		
OX	NCB1_TaxID=1317;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MFE28;		
RX	MEDLINE=90316665; PubMed=2142479;		
RA	Gilmore K.S., Russell R.R., Perretti J.U.;		
RT	"Analysis of the Streptococcus downei gtf5 gene, which specifies a		
RT	glucosyltransferase that synthesizes soluble glucans.";		
RL	Infect. Immun. 58:2452-2458(1990).		
CC	- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT		
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE		
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE		
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.		
CC	- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-		
CC	fructose + {(1,6)-alpha-D-glucosyl} (N+1).		
CC	- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF		
CC	PRIMER GLUCAN UNLIKE GTF-I.		
CC	- DISASE: DENTAL CARIES.		
CC	- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA		
CC	1,6-GLUCOSE).		
CC	- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-		
CC	- BINDING PROTEIN FROM S. MUTANS.		
CC	- SIMILARITY: Contains 10 cell wall binding repeats.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics institute. There are no restrictions on how		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license@sib-sw.ch). or send an email to license@sib-sw.ch).		
CC	EMBL: M30943; AAA26898.1; "		
DR	InterPro: IPR002479; CW_binding.		

DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding 1; 8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 KW TRANSFERRASE; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT CHAIN 1 36
 FT REPEAT 157 177
 FT REPEAT 178 197
 FT REPEAT 198 1061
 FT REPEAT 1062 1082
 FT REPEAT 1083 1102
 FT REPEAT 1150 1169
 FT REPEAT 1170 1190
 FT REPEAT 1225 1243
 FT REPEAT 1289 1308
 FT REPEAT 1309 1328
 FT REPEAT 1331 1352
 FT REPEAT 1365 AA; 151590 MW; 1672965A2E6C476 CRC64;
 SQ SEQUENCE

Query Match 56.3%; Score 67; DB 1; Length 1365;
 Best Local Similarity 57.1%; Pred. No. 0.018;
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANNHSIVEAMSDNDTPYLDH 21
 Db 467 AIDHSLTEAMSGNDNDYKDD 487

RESULT 7
 ID UL16_EBV STANDARD; PRT; 336 AA.
 AC P03221;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE BGLF2 protein.
 DE BGLF2.
 OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
 OS Viruses; daDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Lymphocryptovirus.
 OC NCBI_TaxID=10377;
 OK [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=84270667; PubMed=6087149;
 RX Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
 RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
 RA Tuffnell P.S., Barrett B.G.,
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL Nature 310:207-211 (1984).
 RN [2]
 RP MEDLINE=92113548; PubMed=1662696;
 RX Chen M.R., Hsu T.Y., Lin S.W., Chen J.Y., Yang C.S.,
 RT "Cloning and characterization of cDNA clones corresponding to
 RT transcripts from the BamHI G region of the Epstein-Barr virus genome
 RT and expression of BGLF2.";
 RL J. Gen. Virol. 72:3047-3055 (1991).
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
 CC HSV-6 ORF1R, EBV-1 46, HCW UL94, EBV BDLF2, AND VZV 44.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M60514; AAA45871.1; -
 CC EMBL: V01555; CAA24831.1; -
 CC EMBL: S77132; AAB21113.1; -
 CC PIR: C43044; Q0BEA0.
 DR InterPro: IPR004286; UL16_UL94.

DR Pfam: PF03044; UL16_UL94; 1.
 KW Late protein.
 SQ SEQUENCE 336 AA; 36888 MW; 840937A416D5584C CRC64;
 Query Match 41.2%; Score 49; DB 1; Length 336;
 Best Local Similarity 41.2%; Pred. No. 2.5;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 ANNHSIVEAMSDNDTP 17
 Db 216 AGAHVNLKGTEDDPSF 232

RESULT 8
 ID G3PC LEIME STANDARD; PRT; 330 AA.
 AC Q01558;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
 DE (GAPDH).
 GN GAPC.
 OS Leishmania mexicana.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC NCBI_TaxID=5665;
 OK [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Sep. Mexicana;
 RX MEDLINE=93063042; PubMed=1435864;
 RA Hanaert V., Blaauw M., Kohl L., Allert S., Oppendoerfer F.,
 RA Michels P.A.M.,
 RT "Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-
 RT phosphate dehydrogenase in Leishmania mexicana.";
 RL Mol. Biochem. Parasitol. 55:115-126 (1992).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X65220; CAA46323.1; -
 CC PIR: B48445; B48445.
 DR HSBP; P06977; IGAD.
 DR InterPro: IPR000173; GAP dehydrogenase.
 DR InterPro: IPR006424; GAPDH-1.
 DR Pfam: PF00044; spdh; 1.
 DR Pfam: PF02800; spdh C; 1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR TIGRFS: TIGR01534; GAPDH-1; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 KM Glycolysis; Oxidoreductase; NAD.
 FT INIT MET 0
 FT BINDING 148 148
 FT ACT SITE 175 175
 FT ACT SITE 175 175
 FT ACT SITE 175 175
 SQ SEQUENCE 330 AA; 35511 MW; EDAB6DBE8A207F1E CRC64;
 Query Match 40.3%; Score 48; DB 1; Length 330;
 Best Local Similarity 42.1%; Pred. No. 3.5;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 NNHSIVEAMSDNDTPYLDH 20

DB 300 NDHFVKLVSWYDNETSYSH 318

RESULT 9

YMS8_YEAST STANDARD; PRT; 313 AA.

AC 003695;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 35.0 kDa protein in PER2-HFAl intergenic region.
GN YMR206W OR YMR325.07.
OS *Saccharomyces cerevisiae* (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;
OC *Saccharomycetales*; *Saccharomycetaceae*; *Saccharomyces*.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagers K., Lyne G., Moule S., Odell C., Pearson D., Rajadream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome
XIII".
RL Nature 387:90-93 (1997).
CC -1- SIMILARITY: SOME, TO YEAST YMR014W.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z48755; CAA8648.1; -;
DR PIR; S59448; S59448.
DR SGD; S0004819; YMR206W.
KW Hypothetical protein.
FT DOMAIN 3
FT DOMAIN 146 149 POLY-SER.
FT DOMAIN 246 252 POLY-GLN.
FT DOMAIN 313 AA; 35017 MW; 9D92BDE98257780 CRC64;
SQ SEQUENCE

Query Match 38.7%; Score 46; DB 1; Length 313;
Best Local Similarity 47.4%; Pred. No. 6.9;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 2 NNHVSIVAMSDNDTPYLH 20
DB 125 NNHQTARFRSDSHPSLVH 143

RESULT 10
G3P MYCGE STANDARD; PRT; 337 AA.

AC P47543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP3 OR GAP3 OR MG301.
OS *Mycoplasma genitalium*.
OC Bacteria; Firmicutes; Mollicutes; *Mycoplasmataceae*; *Mycoplasma*.
OX NCBI_TaxID=2097;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

RA Ertchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uitterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of *Mycoplasma genitalium*".
RL Science 270:397-403 (1995).

RN (2)
RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III,
RT "A survey of the *Mycoplasma genitalium* genome by using random
RT sequencing".
RL J. Bacteriol. 175:7918-7930 (1993).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U39710; AAC71523.1; -;
CC EMBL; U02213; AAD12507.1; -;
CC EMBL; U02178; AAD12463.1; -;
CC PIR; C64233; C64233.
CC HSSP; P17721; 1HDG.
CC TIGR; MG301; -;
CC Interpro: IPR000173; GAP dhndrogenase.
CC Interpro: IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh_1.
DR Pfam; PF02800; gpdh_C_1.
DR PRINTS; PR00078; G3PDHDSGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-1; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS
FT FT FT (BY SIMILARITY).
SQ SEQUENCE 337 AA; 37097 MW; FA1EX1966687006B CRC64;

Query Match 37.8%; Score 45; DB 1; Length 337;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 8 YEAMSDNDTPYLH 20
DB 313 VYAMYNDSYVH 325

RESULT 11
SLIF_MOUSE STANDARD; PRT; 569 AA.

AC O92063;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sialic acid binding Ig-like lectin-F precursor (msiglec-F).
GN SIGLEC-F.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)

SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Angata T., Hingorani R., Varki N.M., Varki A.;
 RT "Cloning and characterization of a novel mouse Siglec, mSiglec-F:
 differential evolution of the mouse and human (CD33) Siglec-3-related
 gene clusters.";
 RL J. Biol. Chem. 276:45128-45136(2001).
 CC -1- FUNCTION: Putative adhesion molecule that mediates sialic-acid
 dependent binding to cells. Preferentially binds to alpha2,3-
 linked sialic acid. The sialic acid recognition site may be masked
 by cis interactions with sialic acids on the same cell surface.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Predominantly expressed by immature
 monocytic/myeloid lineage cells in bone marrow. Also found at
 lower levels in mature neutrophils and monocytes.
 CC -1- DOMAIN: Contains an intracytoplasmic motif referred as
 immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif
 is involved in downmodulation of cellular functions as the
 termination of the immune response.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
 (SIALIC ACID BINDING Ig-LIKE LECTIN) SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL: AF293371; AAL1043.1; -
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig; 2.
 DR SMART: SM00408; IGC2; 1.
 DR PROSITE: PS50835; IG_LIKE; 2.
 DR Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 16
 FT CHAIN 17 569
 FT DOMAIN 17 439
 FT TRANSMEM 440 460
 FT DOMAIN 461 569
 FT DOMAIN 18 116
 FT DOMAIN 139 224
 FT DOMAIN 229 324
 FT SITE 536 541
 FT SITE 559 564
 FT SITE 35 163
 FT DISULFID 35 163
 FT DISULFID 40 96
 FT DISULFID 157 206
 FT DISULFID 265 308
 FT CARBOHYD 95 95
 FT CARBOHYD 151 151
 FT CARBOHYD 200 200
 FT CARBOHYD 203 203
 FT CARBOHYD 369 369
 FT CARBOHYD 372 372
 FT CARBOHYD 387 387
 SQ SEQUENCE 569 AA; 61476 MW; 8093838090484FC1 CRC64;
 Query Match 37.0%; Score 44; DB 1; Length 569;
 Best Local Similarity 57.1%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

RESULT 12
 YG3C YEAST STANDARD; PRT; 1122 AA.
 ID YG3C YEAST
 AC P53280;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 128.8 kDa protein in PAS2-PRS5 intergenic region.
 GN YGR134W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Talia E.,
 RA Nawrocki A., del Bino S., Goffeau A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL: Z72919; CA97147.1; -
 DR PIR: S6443; S6443.
 DR SCD: S0003366; CAF130.
 DR GO: GO:0030014; C:CCR4-NOT complex; IPI.
 DR GO: GO:0006357; P:regulation of transcription from pol II pro. .; IPI.
 DR KW Hypothetical protein.
 SQ SEQUENCE 1122 AA; 128779 MW; BF3542816CC78490 CRC64;
 Query Match 37.0%; Score 44; DB 1; Length 1122;
 Best Local Similarity 46.7%; Pred. No. 68;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

1 ANNHVSIVEAMSDND 15
 Db 110 SSVASYLSWKDND 124
 RESULT 13
 YPO4 CAEEL STANDARD; PRT; 1131 AA.
 ID YPO4 CAEEL
 AC P53015;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 127.2 kDa protein F31E3.4 in chromosome III.
 GN F31E3.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Favello A.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO YEAST PAN2 AND TO S.POMBE SPAC2267.04.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

```

DR EMBL: U40935; AAA81687.1; -
DR PIR: T16217; T16217.
DR WormPep: F31E3.4; CE01269.
DR InterPro: IPR006055; Exonuclease.
DR Pfam: PF00929; Exonuclease; 1.
DR SMART: SM00479; EXOIII; 1.
DR Hypothetical protein.
SQ SEQUENCE 1131 AA; 127174 MW; D5544AB0C309E7DE CRC64;

Query Match 37.0%; Score 44; DB 1; Length 1131;
Best Local Similarity 36.8%; Pred. No. 69;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 NH5IVEAVMSDNDTPYLHD 21
Db 315 NHGIVNVFADRDPOVNE 333

RESULT 14
PGCV_MOUSE STANDARD; PRT; 3358 AA.
ID PGCV_MOUSE STANDARD; PRT; 3358 AA.
AC Q62059; Q62058; Q9CTU0;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSFG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shimomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of Pg-M(V3), an alternatively spliced form of Pg-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues."
RT J. Biol. Chem. 270:3914-3918(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM V3).
RC STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shimomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of Pg-M(V3), an alternatively spliced form of Pg-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues."
RT J. Biol. Chem. 270:3914-3918(1995).
RN [3]
RP SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Komuro H., Adachi T., Fukuda S.,
RA Arizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann M., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai T., Pesole G., Quackenbush J.,
RA Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga M., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guselnich S., Hill D., Hornum M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyono O.K., Wang K.H., Welter C., Whitlaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).

```

```

RN [4]
RP INTERACTION WITH FBIN1.
RX PubMed=10400671;
RA Asberg A., Adam S., Kostka G., Timpel R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican."
RT J. Biol. Chem. 274:20444-20449(1999).
CC - FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC - SUBUNIT: Interacts with FBIN1.
CC - SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC - ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087; VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087; VSP_003090;
CC - TISSUE SPECIFICITY: V2 is found only in brain.
CC - DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC - SIMILARITY: Contains 2 EGF-like domains.
CC - SIMILARITY: Contains 2 EGF-like domains.
CC - SIMILARITY: Contains 1 C-type lectin family domain.
CC - SIMILARITY: Contains 1 Sushi (SCR) domain.
CC - SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: D16263; BAA03796.1; -
DR EMBL: D26599; -; NOT ANNOTATED_CDS.
DR EMBL: D32040; BAA06802.1; -
DR EMBL: AK014525; BAB29411.1; -
DR HSSP: P01132; IEPG.
DR MGD: MGI:102889; Cspg2.
DR InterPro: IPR000152; Aex_hydroxyl.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR007110; I9-like.
DR InterPro: IPR003599; I9.
DR InterPro: IPR001306; I9_MHC.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR000538; Link.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00047; I9; 1.
DR Pfam: PF00059; Lectin_C; 1.
DR Pfam: PF00084; sushi; 1.
DR Pfam: PF00193; Xlink; 2.
DR PRINTS: PR01265; LINKMODULE.
DR ProDom: PD000918; Link; 2.
DR SMART: SM00032; CCP; 1.
DR SMART: SM00034; CLEC7; 1.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00409; I9; 1.
DR SMART: SM00445; LINK; 2.
DR SMART: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE: PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; 2.

```

DR PROSITE; PS01186; EGF 2; 1.
 DR PROSITE; PS01187; EGF CA; 1.
 DR PROSITE; PS00835; IG Like; 1.
 DR PROSITE; PS01241; LINK; 2.
 KM Glycoprotein; Lectin; Extracellular matrix; Sushi;
 KM Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KM Hyaluronic acid; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 3358
 FT DOMAIN 167 146
 FT DOMAIN 167 244
 FT DOMAIN 265 346
 FT DOMAIN 348 1308
 FT DOMAIN 1309 3052
 FT DOMAIN 3052 3088
 FT DOMAIN 3090 3126
 FT DOMAIN 3139 3253
 FT DOMAIN 3258 3316
 FT DISULFID 44 130
 FT DISULFID 172 243
 FT DISULFID 196 217
 FT DISULFID 270 333
 FT DISULFID 294 315
 FT DISULFID 3056 3076
 FT DISULFID 3061 3076
 FT DISULFID 3078 3087
 FT DISULFID 3094 3105
 FT DISULFID 3099 3114
 FT DISULFID 3116 3125
 FT DISULFID 3132 3143
 FT DISULFID 3160 3252
 FT DISULFID 3228 3244
 FT DISULFID 3259 3302
 FT DISULFID 3288 3315
 FT CARBOHYD 57 57
 FT CARBOHYD 330 330
 FT CARBOHYD 351 351
 FT CARBOHYD 441 441
 FT CARBOHYD 807 807
 FT CARBOHYD 914 914
 FT CARBOHYD 951 951
 FT CARBOHYD 1305 1305
 FT CARBOHYD 1372 1372
 FT CARBOHYD 1679 1679
 FT CARBOHYD 2054 2054
 FT CARBOHYD 2244 2244
 FT CARBOHYD 2362 2362
 FT CARBOHYD 2627 2627
 FT CARBOHYD 3030 3030
 FT CARBOHYD 3332 3332
 FT CARBOHYD 3342 3342
 FT CARBOHYD 348 348
 FT VARSPPLIC 349 1308
 FT VARSPPLIC 1309 3052
 FT VARSPPLIC 349 3052
 FT CONFLICT 126 126
 FT CONFLICT 348 348
 FT CONFLICT 1658 1658
 FT CONFLICT 1674 1680
 FT SEQUENCE 3358 AA; 366938 MW; 07180026BC0762D CRC64;

Query Match 37.0%; Score 44; DB 1; Length 3358;
 Best Local Similarity 50.0%; Fred. No. 2.6e+02;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 NHVSIKAWSDNDPRY 18
 DB 510 NHISLKLKPKKNTY 525

RESULT 15
 ID G6PD_ECOLI STANDARD: PRT: 491 AA.
 AC P22992; P78069; Q60134; Q60139;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (G6PD).
 GN ZWF OR B1852 OR C2265.
 OS Escherichia coli, and
 OS Escherichia coli 06.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562, 217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=91123224; PubMed=1704005;
 RT Rowley D.L., Wolf R.E. Jr.;
 RL "Molecular characterization of the Escherichia coli K-12 zwf gene encoding glucose 6-phosphate dehydrogenase.";
 RN J. Bacteriol. 173:968-977(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1234-1238(1997).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Maki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshino H.,
 RA Saito N., Samped J., Seki Y., Sivasubramanian S., Tagami H.,
 RA Takekida J., Takemoto K., Mada C., Yamamoto Y., Horikuchi T.;
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=2288234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 [5]
 RP SEQUENCE OF 72-368 FROM N.A.
 RC STRAIN=Various ECOR strains;
 RX MEDLINE=95064015; PubMed=7973728;
 RA Gutman D.S., Dykhuizen D.E.;
 RT "Clonal divergence in Escherichia coli as a result of recombination, not mutation.";
 RL Science 266:1380-1383(1994).
 [6]
 RP SEQUENCE OF 321-491 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93345818; PubMed=8344525;
 RA Carter A.T., Pearson B.M., Dickinson J.R., Larcaschire W.E.;
 RT "Sequence of the Escherichia coli K-12 edd and eda genes of the Emmer-Doudoroff pathway.";

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:31:40 / Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	119	100.0	1590	2	Q59983 streptococc
2	119	100.0	1590	2	Q55263 streptococc
3	76	63.9	1016	2	Q91CJ7 leuconostoc
4	72	60.5	2835	2	Q8G9Q2 leuconostoc
5	68	57.1	1477	2	Q91466 leuconostoc
6	68	57.1	1508	2	Q9E2H5 leuconostoc
7	68	57.1	1508	2	Q52224 leuconostoc
8	68	57.1	1575	2	Q91CH3 streptococc
9	68	57.1	1577	2	Q54178 streptococc
10	67	56.3	1338	2	Q9WXJ4 streptococc
11	65	54.6	1527	2	Q9ZAR4 leuconostoc
12	65	54.6	1527	2	Q8KRE1 leuconostoc
13	61	51.3	1449	2	Q68542 streptococc
14	61	51.3	1449	2	Q55264 streptococc
15	61	51.3	1554	2	Q8KZL5 streptococc
16	60	50.4	1512	2	Q9WXJ5 streptococc

17	57	47.9	1577	2	Q55265 streptococc
18	56	47.1	1518	2	Q00600 streptococc
19	55	46.2	1599	2	Q00599 streptococc
20	55	46.2	2057	2	Q9RE05 leuconostoc
21	53	44.5	175	10	Q91MJ8 arabidopsis
22	53	44.5	195	10	Q8LDM8 arabidopsis
23	51	42.9	1195	2	Q06459 thermus the
24	50	42.0	601	5	Q9VOR4 drosophila
25	50	42.0	630	5	Q81FA3 drosophila
26	49	41.2	524	16	Q9K318 drosophila
27	48.5	40.8	282	10	Q8RWJ3 drosophila
28	48	40.3	184	16	Q8XME4 drosophila
29	48	40.3	380	3	Q8N773 drosophila
30	47	39.5	193	2	Q9S445 drosophila
31	46.5	39.1	212	2	Q9K167 drosophila
32	46.5	39.1	214	11	Q8C1P7 drosophila
33	46.5	39.1	1604	4	Q8NPA0 drosophila
34	46	38.7	169	5	Q8S589 drosophila
35	46	38.7	271	11	Q9QZEE drosophila
36	46	38.7	283	5	Q8MSF4 drosophila
37	46	38.7	283	5	Q8Y2Q6 drosophila
38	46	38.7	393	16	Q8Y2Q6 drosophila
39	46	38.7	490	16	Q06579 mycobacteri
40	46	38.7	601	16	Q05566 mycobacteri
41	46	38.7	810	5	Q8T3J2 drosophila
42	46	38.7	811	5	Q9VK54 drosophila
43	46	38.7	880	5	P91643 drosophila
44	45.5	38.2	418	17	Q97U15 sulfolobus
45	45	37.8	119	2	Q93RH1 streptococc

ALIGNMENTS

RESULT 1

059983 PRELIMINARY; PRT; 1590 AA.

AC 059983; 01-NOV-1996 (TREMUR1.01, Created)
DT 01-NOV-1996 (TREMUR1.01, Last sequence update)
DT 01-OCT-2002 (TREMUR1.22, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5).
GN GTFI.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OM2176;
RA MEDLINE=94146405; PubMed=8312602;
RX Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Karayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
Streptococcus sobrinus.";
RL DNA Seq. 4:19-27(1993).
DR EMBL, D13858; BA002976.1; -;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glycosyltransferase; Signal; Transferase.
FT SIGNAL 1
FT CHAIN 39 1590 GLUCOSYLTRANSFERASE-I.
SQ SEQUENCE 1590 AA; 17595 MW; C3C83A57CF3C280E CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAWSNDTPYLHD 21
DB 477 ANNHVSIVEAWSNDTPYLHD 497

RESULT 2

Q55263 PRELIMINARY; PRT; 1590 AA.
 AC Q55263;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478.";
 RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 15.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
 ID 477 ANNHVSIVEAWSNDNTPYLHD 497

RESULT 3
 Q91CJ7 PRELIMINARY; PRT; 1016 AA.
 AC Q91CJ7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DSRT.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 CX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512F;
 RX MEDLINE=20169623; PubMed=10705445; Kobayashi M.;
 RA Funane K., Mizuno K., Takahara H.;
 RT "Gene encoding a dextranucrase-like protein in Leuconostoc mesenteroides NRRL B-512F.";
 RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
 DR EMBL; AB020020; BAA90527.1; -;
 DR HSBP; P06278; 1VJ5.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1016 AA; 110344 MW; 8896EFDE13CCBA7 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
 Best Local Similarity 61.9%; Pred. No. 0.0042;
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
 ID 548 ANNHVSIVEAWSNDNTPYLHD 568

RESULT 4

Q8G9Q2 PRELIMINARY; PRT; 2835 AA.
 AC Q8G9Q2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5) (Fragment).
 GN DSRE.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 CX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2231661; PubMed=12270834;
 RA Bozonnet S., Dols-Lafargue M., Fabre E., Pizzut S., Renaud-Simeon M.,
 RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
 RT synthesizing dextranucrase with two catalytic domains.";
 RL J. Bacteriol. 184:5753-5761(2002).
 DR EMBL; AJ430204; CAD22883.1; -;
 DR Transferase; Glycosyltransferase.
 KW NON TER
 FT 1
 SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match 60.5%; Score 72; DB 2; Length 2835;
 Best Local Similarity 57.9%; Pred. No. 0.058;
 Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYL 19
 ID 557 ANNHVSIVEAWSNDNTPYL 575

RESULT 5
 Q91A66 PRELIMINARY; PRT; 1477 AA.
 AC Q91A66;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRC.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 CX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon M., Pizzut S., Sarcabal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
 RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBS databases.
 DR EMBL; AJ250172; CAB76565.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Glycosyltransferase; Transferase.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710BDFCB831 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1477;
 Best Local Similarity 57.1%; Pred. No. 0.12;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
 ID 532 ANNHVSIVEAWSNDNTPYLHD 552

RESULT 6
 Q9EZHS PRELIMINARY; PRT; 1508 AA.
 ID Q9EZHS

```

AC Q95ZH5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL PubMed=14932000 to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAC38021.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87A8AF3A CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLMD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 7
ID 052224 PRELIMINARY; PRT; 1508 AA.
AC 052224;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL EMBL; AF030129; AAB95453.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLMD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 8
ID 091CH3 PRELIMINARY; PRT; 1575 AA.
AC 091CH3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RA MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Oshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483 (2000).
DR EMBL; AB025228; BAA95201.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2E543 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1575;
Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLMD 21
Db 546 AIRHLSILEAWSNDPEYKD 566

RESULT 9
ID 054178 PRELIMINARY; PRT; 1577 AA.
AC 054178; Q54247;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RA MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314 (1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RC MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgg, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586 (1992).
DR EMBL; U12643; AAC43483.1; -
DR EMBL; M89776; AAB26969.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase.
SQ SEQUENCE 1577 AA; 177805 MW; 5AF0328DC5E08D18 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1577;

```


Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 548 ALKHLSTLEAWSNDPPLYHD 568

RESULT 10

ID Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
AC Q9WXJ4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF-S.
OS Streptococcus criceti.
OC Plasmid pAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S. criceti glucosyltransferase (glfs and gftf) genes."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB026123; BAA7236.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E1SD99B CRC64;

Query Match 56.3%; Score 67; DB 2; Length 1338;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 437 AIDHLSLEAWSNDPPLYHD 457

RESULT 11
ID Q9ZAR4 PRELIMINARY; PRT; 1527 AA.
AC Q9ZAR4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase.
GN DEX.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512-F;
RA Bhattacharjee R., Singh D.K.S.;
RT "Cloning and Molecular Characterization of Dextranucrase Gene from
Leuconostoc mesenteroides NRRL B-512-F."
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; U81374; AAD10952.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1527 AA; 169709 MW; 1DEFAFA237C743398 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 581 ANQHLSLEAWSNDPPLYHD 601

RESULT 12

ID Q8KRE1 PRELIMINARY; PRT; 1527 AA.
AC Q8KRE1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Dextranucrase Dsrd (BC 2.4.1.5).
GN DSRD.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RA Neubauer H., Hauche A., Mollet B.;
RT "Isolation and characterization of the dextranucrase Dsrd of
Leuconostoc mesenteroides Lcc4."
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY017384; AAG6158.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferrase; Glycosyltransferase.
KM Transferrase.
SQ SEQUENCE 1527 AA; 169835 MW; F9D0DE2205D893668 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 581 ANQHLSLEAWSNDPPLYHD 601

RESULT 13
ID Q68542 PRELIMINARY; PRT; 1449 AA.
AC Q68542;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VLA477;
RA Jaffe R.I.;
RT "Streptococcus salivarius VLA477 gtfN."
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferrase.
KT NON TER
FT NON TER
SQ SEQUENCE 1449 AA; 159895 MW; 0700FED748471BFB CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAMSDNDTPYLHD 21
 DB 536 AIKHLSILEAMSHNDAYNED 556

RESULT 14
 ID 055264 PRELIMINARY; PRT; 1449 AA.
 AC 055264;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RL Infect. Immun. 63:609-621(1995).
 DR EMBL; L35495; AAC41412.1; -
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Signal; transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MW; D62F07306E86A46 CRC64;
 Query Match 51.3%; Score 61; DB 2; Length 1449;
 Best Local Similarity 57.1%; Pred. No. 1.4;
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAMSDNDTPYLHD 21
 DB 536 AIKHLSILEAMSHNDAYNED 556

RESULT 15
 ID 08KZL5 PRELIMINARY; PRT; 1554 AA.
 AC 08KZL5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTF.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sengoku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan.";
 RL Biochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL; AB089438; BAC07265.1; -
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM transferase.

SQ SEQUENCE 1554 AA; 171676 MW; 6981BC1DAE24A73 CRC64;
 Query Match 51.3%; Score 61; DB 2; Length 1554;
 Best Local Similarity 61.1%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 HVSIVEAMSDNDTPYLHD 21
 DB 488 HISILEAMSLINDQYNED 505

Search completed: November 13, 2003, 09:44:04
 Job time : 32.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVFAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCPUS COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	90.8	1475	3	US-09-007-999-2
2	108	90.8	1475	3	US-09-210-361-2
3	108	90.8	1475	3	US-09-740-274-2
4	99	83.2	1375	3	US-09-210-361-4
5	99	83.2	1375	4	US-09-740-274-4
6	74	62.2	1430	3	US-09-604-957-4
7	74	62.2	1430	3	US-09-008-172-2
8	74	62.2	1430	3	US-09-210-361-6
9	74	62.2	1430	3	US-09-740-274-6
10	65	54.6	523	4	US-09-604-957-5
11	57	47.9	1577	2	US-08-793-824-2
12	55	46.2	584	4	US-09-604-957-6
13	55	46.2	2057	4	US-09-499-203-2
14	48	40.3	535	4	US-09-604-957-7
15	48	40.3	1278	4	US-09-604-957-3
16	47	39.5	749	4	US-09-562-737-96
17	43.5	36.6	1092	4	US-09-275-608-3
18	43	36.1	151	4	US-09-198-452A-815
19	43	36.1	234	4	US-09-134-001C-4514
20	43	36.1	647	4	US-09-134-001C-5458
21	42	35.3	200	2	US-08-836-442-4
22	42	35.3	287	4	US-09-252-991A-31548
23	42	35.3	327	4	US-09-107-532A-6181
24	42	35.3	334	6	5290690-11
25	42	35.3	403	4	US-09-252-991A-27960
26	42	35.3	454	4	US-09-252-991A-28000
27	42	35.3	455	3	US-09-362-473-6

28	42	35.3	529	2	US-08-836-442-3	Sequence 3, Appli
29	42	35.3	1627	1	US-07-665-792E-9	Sequence 9, Appli
30	42	35.3	3898	3	US-08-750-717-2	Sequence 2, Appli
31	41.5	34.9	69	4	US-09-252-991A-17383	Sequence 17383, A
32	41.5	34.9	130	4	US-09-328-352-4952	Sequence 4952, Ap
33	41	34.5	88	4	US-09-252-991A-21480	Sequence 21480, A
34	41	34.5	259	1	US-08-277-231A-3	Sequence 6, Appli
35	41	34.5	259	2	US-08-473-750-6	Sequence 6, Appli
36	41	34.5	259	2	US-08-477-326-6	Sequence 3, Appli
37	41	34.5	2465	2	US-08-596-291-3	Sequence 3, Appli
38	41	34.5	2465	3	US-09-100-804-3	Sequence 12, Appli
39	41	34.5	2466	4	US-09-080-855-12	Sequence 12, Appli
40	41	34.5	2466	4	US-09-566-076-12	Sequence 2, Appli
41	41	34.5	2466	5	PCT-US94-09943-2	Sequence 46, Appli
42	41	34.5	2485	3	US-09-290-640-46	Sequence 2, Appli
43	41	34.5	3898	2	US-08-876-991-2	Sequence 2, Appli
44	41	34.5	3898	2	US-09-059-853-2	Sequence 2, Appli
45	40.5	34.0	214	3	US-09-214-278-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ. ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ. ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANNHVSIVFAMSDNDTPYLHD 21
DB 481 ANNHVSIVFAMSDNDTPYLHD 501

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

```

```

Query Match          90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      481 ANDHSLTEAMSDNDTPYLHD 501

```

```

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

```

```

Query Match          90.8%; Score 108; DB 4; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      481 ANDHSLTEAMSDNDTPYLHD 501

```

```

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

```

```

Query Match          83.2%; Score 99; DB 3; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      507 ANDHSLTEAMSYNDTPYLHD 527

```

```

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

```

```

Query Match          83.2%; Score 99; DB 4; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      507 ANDHSLTEAMSYNDTPYLHD 527

```

```

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314

```

```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 4338
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-604-957-4

```

```

Query Match          62.2%; Score 74; DB 4; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.00098;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 75 AINHLSILEAWSNDPQYNKD 95

```

```

RESULT 7
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-008-172-2

```

```

Query Match          62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 495 AINHLSILEAWSNDPQYNKD 515

```

```

RESULT 8
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07

```

```

; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-210-361-6

```

```

Query Match          62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 495 AINHLSILEAWSNDPQYNKD 515

```

```

RESULT 9
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-6

```

```

Query Match          62.2%; Score 74; DB 4; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 495 AINHLSILEAWSNDPQYNKD 515

```

```

RESULT 10
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT

```

APPLICANT: RAHAOUI, HAKIM
 APPLICANT: LEER, ROBERT-JAN
 TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
 FILE REFERENCE: BO 43388
 CURRENT APPLICATION NUMBER: US/09/604,957
 CURRENT FILING DATE: 2000-06-28
 PRIOR APPLICATION NUMBER: 00201871.1
 PRIOR FILING DATE: 2000-05-25
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 5
 LENGTH: 523
 TYPE: PRT
 ORGANISM: Leuconostoc mesenteroides
 US-09-604-957-5

Query Match 54.6%; Score 65; DB 4; Length 523;
 Best Local Similarity 57.1%; Pred. No. 0.025;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANHVSIVAMSDNDTPYLHD 21
 Db 75 ANQHLSILEDWNGKDPQYVN 95

RESULT 11
 US-08-793-824-2
 Sequence 2, Application US/08793824

GENERAL INFORMATION:
 APPLICANT: Simpson, Christine Lynn
 APPLICANT: Giffard, Philip Morrison
 APPLICANT: Jacques, Nicholas Anthony
 TITLE OF INVENTION: Genetic Manipulation of Plants to
 INCREASE STARCH CONTENT
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Griffith Hack & Co
 STREET: Level 8, 168 Walker Street
 CITY: No. 5981838th Sydney
 STATE: New South Wales
 COUNTRY: Australia
 ZIP: 2060

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/793,824
 FILING DATE:
 CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PM7643

FILING DATE: 24-AUG-1994

TELECOMMUNICATION INFORMATION:

TELEPHONE: 61 2 9957 5944

TELEFAX: 61 2 957 6288

TELEX: 26547

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1577 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: not relevant

MOLECULE TYPE: Protein

ORIGINAL SOURCE:

ORGANISM: Streptococcus salivarius

Query Match 47.9%; Score 57; DB 2; Length 1577;
 Best Local Similarity 61.1%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 HVSIVAMSDNDTPYLHD 21
 Db 591 HSLSEAMSYNDHQYKND 608

RESULT 12

US-09-604-957-6

Sequence 6, Application US/09604957

Patent No. 6486314

GENERAL INFORMATION:

APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA

APPLICANT: DIJKHUIZEN, IUBERT

APPLICANT: RAHAOUI, HAKIM

APPLICANT: LEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN

FILE REFERENCE: BO 43388

CURRENT APPLICATION NUMBER: US/09/604,957

CURRENT FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 00201871.1

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 6

LENGTH: 584

TYPE: PRT

ORGANISM: Leuconostoc mesenteroides

US-09-604-957-6

Query Match 46.2%; Score 55; DB 4; Length 584;
 Best Local Similarity 45.0%; Pred. No. 1.1;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANHVSIVAMSDNDTPYLH 20
 Db 75 ANKHLSILEDWNGKDPQYVN 94

RESULT 13

US-09-499-203-2

Sequence 2, Application US/09499203

Patent No. 6570065

GENERAL INFORMATION:

APPLICANT: KOSSMANN, Jens

APPLICANT: WELSH, Thomas

APPLICANT: QUANZ, Martin

APPLICANT: KNUTH, Karola

TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase

FILE REFERENCE: 147-196P

CURRENT APPLICATION NUMBER: US/09/499,203

CURRENT FILING DATE: 2000-02-08

NUMBER OF SEQ ID NOS: 54

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 2

LENGTH: 2057

TYPE: PRT

ORGANISM: Leuconostoc mesenteroides

US-09-499-203-2

Query Match 46.2%; Score 55; DB 4; Length 2057;
 Best Local Similarity 45.0%; Pred. No. 4.5;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANHVSIVAMSDNDTPYLH 20
 Db 665 ANKHLSILEDWNGKDPQYVN 684

RESULT 14

US-09-604-957-7

Sequence 7, Application US/09604957

Patent No. 6486314

GENERAL INFORMATION:

```

; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

```

Query Match          40.3%; Score 48; DB 4; Length 535;
Best Local Similarity 35.0%; Pred. No. 12;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      1 ANHWSTVEAMSDNDTPEYLH 20
DB      74 SNKHINILEDWMMHADPEYFN 93

```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

```

```

Query Match          40.3%; Score 48; DB 4; Length 1278;
Best Local Similarity 35.0%; Pred. No. 34;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      1 ANHWSTVEAMSDNDTPEYLH 20
DB      550 SNKHINILEDWMMHADPEYFN 569

```

```

Search completed: November 13, 2003, 09:48:00
Job time : 14.6351 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	90.8	1475	US-09-740-274-2	Sequence 2, Appli
2	99	83.2	1375	US-09-740-274-4	Sequence 4, Appli
3	74	62.2	545	US-09-995-749A-10	Sequence 6, Appli
4	74	62.2	1430	US-09-740-274-6	Sequence 11, Appli
5	65	54.6	522	US-09-995-749A-11	Sequence 12, Appli
6	55	46.2	584	US-09-995-749A-12	Sequence 13, Appli
7	48	40.3	535	US-09-995-749A-13	Sequence 2, Appli
8	48	40.3	1781	US-09-995-749A-2	Sequence 96, Appli
9	47	39.5	749	US-10-211-962-96	Sequence 73, Appli
10	46.5	39.1	1604	US-09-888-615-73	Sequence 1, Appli
11	46	38.7	502	US-10-009-823A-1	Sequence 17, Appli
12	46	38.7	1494	US-10-195-144-17	Sequence 4571, Ap
13	45	37.8	199	US-09-738-626-4571	Sequence 14923, A
14	45	37.8	665	US-10-156-761-14923	Sequence 16, Appli
15	45	37.8	972	US-09-924-154-16	

15	44	37.0	15	12	US-10-295-693-90	Sequence 90, Appli
17	44	37.0	15	15	US-10-067-649-90	Sequence 90, Appli
18	43.5	36.6	202	10	US-09-280-197-11	Sequence 11, Appli
19	43.5	36.6	202	12	US-10-448-139-11	Sequence 11, Appli
20	43.5	36.6	491	9	US-09-815-242-10190	Sequence 10190, A
21	43.5	36.6	491	9	US-09-815-242-13756	Sequence 13756, A
22	43.5	36.6	1092	10	US-09-423-126-5	Sequence 26, Appli
23	43	36.1	27	15	US-10-092-908-26	Sequence 324, App
24	43	36.1	67	15	US-10-050-704-234	Sequence 149, App
25	43	36.1	87	15	US-10-050-704-149	Sequence 20, Appli
26	43	36.1	332	10	US-09-454-279-20	Sequence 66, Appli
27	43	36.1	437	11	US-09-374-046A-66	Sequence 34, Appli
28	43	36.1	539	13	US-10-230-026-34	Sequence 2, Appli
29	43	36.1	603	15	US-10-054-678-35	Sequence 35, Appli
30	43	36.1	603	15	US-10-092-908-35	Sequence 7562, Ap
31	43	36.1	715	15	US-10-156-761-7562	Sequence 22, Appli
32	42	35.3	349	12	US-10-220-381-22	Sequence 5096, Ap
33	42	35.3	432	9	US-09-815-242-5096	Sequence 6, Appli
34	42	35.3	455	9	US-09-742-954-6	Sequence 28981, A
35	42	35.3	7349	12	US-10-314-657-46	Sequence 146, App
36	41.5	34.9	49	12	US-10-029-386-28981	Sequence 146, App
37	41	34.5	120	14	US-10-001-835-146	Sequence 14004, A
38	41	34.5	302	9	US-09-815-242-14004	Sequence 9, Appli
39	41	34.5	358	11	US-09-095-478-9	Sequence 7, Appli
40	41	34.5	381	11	US-09-095-478-7	Sequence 14114, A
41	41	34.5	438	15	US-10-156-761-14114	Sequence 11729, A
42	41	34.5	749	15	US-10-156-761-11729	Sequence 56, Appli
43	41	34.5	1267	15	US-10-059-585-56	Sequence 12, Appli
44	41	34.5	2466	12	US-10-177-980-12	Sequence 46, Appli
45	41	34.5	2485	9	US-09-802-669-46	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 90.8%; Score 108; DB 9; Length 1475;
Best Local Similarity 85.7%; Pred. No. 7.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
1 ANNHVSIVEAWSNDTPYLHD 21
||:|||||

DB 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 83.2%; Score 99; DB 9; Length 1375;
Best Local Similarity 81.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21

DB 507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 3
US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 62.2%; Score 74; DB 10; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21
DB 75 AINHLSILEAWSNDNTPQYNKD 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 62.2%; Score 74; DB 9; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21

DB 495 AINHLSILEAWSNDNTPQYNKD 515

RESULT 5
US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 54.6%; Score 65; DB 10; Length 522;
Best Local Similarity 57.1%; Pred. No. 0.097;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAMSDNDTPYLH 21
DB 75 ANKHSTLEDMWGNDDPYLV 95

RESULT 6

US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995, 749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604, 957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconobloc mesenteroides
US-09-995-749A-12

Query Match 46.2%; Score 55; DB 10; Length 584;
Best Local Similarity 45.0%; Pred. No. 3.8;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAMSDNDTPYLH 20
DB 75 ANKHSTLEDMWGNDDPYLV 94

RESULT 7

US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995, 749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604, 957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 40.3%; Score 48; DB 10; Length 535;
Best Local Similarity 35.0%; Pred. No. 4.1;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAMSDNDTPYLH 20
DB 74 SNKHITLEDWGNADPEYFN 93

RESULT 8

US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995, 749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604, 957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 40.3%; Score 48; DB 10; Length 1781;
Best Local Similarity 35.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAMSDNDTPYLH 20
DB 1053 SNKHITLEDWGNADPEYFN 1072

RESULT 9

US-10-211-962-96
; Sequence 96, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211, 962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562, 737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-10-211-962-96

Query Match 39.5%; Score 47; DB 15; Length 749;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 6 SIVEAMSDNDTPYLH 20
DB 136 SAVEEMTDSATPHPH 150

RESULT 10

US-09-888-615-73
; Sequence 73, Application US/09888615
; Patent No. US20020064856A1

```

; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHAYDOZAK, GLEN
; APPLICANT: MANNING, GERRARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 1604
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-73

```

```

Query Match          39.1%; Score 46.5; DB 9; Length 1604;
Best Local Similarity 47.4%; Pred. No. 2.4e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
QY 5 VSIVEAMSDN--DTPYLH 20
DB 294 VALLEWVKNDRTDIDPELH 312

```

```

RESULT 11
US-10-009-823A-1
; Sequence 1, Application US/10009823A
; Publication No. US20030157120A1
; GENERAL INFORMATION:
; APPLICANT: Panaccio, Michael
; APPLICANT: Rosey, Everett Lee
; APPLICANT: Sinistaj, Merti
; APPLICANT: Hasse, Detlef
; APPLICANT: Parsons, Jim
; APPLICANT: Ankenbauer, Robert G.
; TITLE OF INVENTION: LAMSONIA DERIVED GENE AND RELATED FLGE
; TITLE OF INVENTION: POLYPEPTIDES, PEPTIDES AND PROTEINS AND THEIR USES
; FILE REFERENCE: DAV1150.001APC
; CURRENT APPLICATION NUMBER: US/10/009,823A
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00437
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/133,973
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-10-009-823A-1

```

```

Query Match          38.7%; Score 46; DB 12; Length 502;
Best Local Similarity 41.2%; Pred. No. 79;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAMSDNDTP 17
DB 185 ANPYFALLESWKNGTGP 201

```

```

RESULT 12
US-10-195-144-17
; Sequence 17, Application US/10195144
; Publication No. US20030126646A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.

```

```

; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENOTT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; FILE REFERENCE: 16313-0136
; CURRENT APPLICATION NUMBER: US/10/195,144
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1494
; TYPE: PRT
; ORGANISM: Rapphanns sativum
US-10-195-144-17

```

```

Query Match          38.7%; Score 46; DB 15; Length 1494;
Best Local Similarity 41.2%; Pred. No. 2.7e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 3 NHVSIVEAMSDNDTPYL 19
DB 1130 NTISLAFTWLONSEPFL 1146

```

```

RESULT 13
US-09-738-626-4571
; Sequence 4571, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4571
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4571

```

```

Query Match          37.8%; Score 45; DB 10; Length 199;
Best Local Similarity 58.3%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 7 IVEAMSDNDTPY 18
DB 177 LVEWTDIDTPY 188

```

RESULT 14

US-10-156-761-14923
 ; Sequence 14923, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 14923
 ; LENGTH: 665
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-14923

Query Match

Best Local Similarity 37.8%; Score 45; DB 15; Length 665;
 Pred. No. 1.5e+02;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 3 NHVSIVEAWSNDPTPYLH 20

DB 88 NMTYLVIEWSDPYTTWRH 105

RESULT 15

US-09-924-154-16
 ; Sequence 16, Application US/09924154
 ; Patent No. US20020127241A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Narum, David L.
 ; APPLICANT: Sim, Kim L.
 ; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
 ; FILE REFERENCE: 05213-0465 43170-262105
 ; CURRENT APPLICATION NUMBER: US/09/924,154
 ; CURRENT FILING DATE: 2001-08-07
 ; PRIOR APPLICATION NUMBER: US 60/223,525
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 972
 ; TYPE: PRT
 ; ORGANISM: Mammalian
 US-09-924-154-16

Query Match

Best Local Similarity 37.8%; Score 45; DB 10; Length 972;
 Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 3 NHVSIVEAWSNDPTPYLH 20

DB 304 NHTGVCLMEDDNDNGYLH 321

Search completed: November 13, 2003, 10:29:03
 Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:11:40 (Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049A-13
Perfect score: 113
Sequence: 1 AIDHLSILEAMSGNDNDYKQ 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	75	66.4	1430 23 AAU98029	S. mutans glucosyl
2	75	66.4	1430 23 AAU98041	S. mutans glucosyl
3	75	66.4	1430 23 AAU98042	S. mutans glucosyl
4	75	66.4	1430 23 AAU98043	S. mutans glucosyl
5	75	66.4	1430 23 AAU98044	S. mutans glucosyl
6	75	66.4	1430 23 AAU98045	S. mutans glucosyl
7	73	64.6	1017 23 AAU79285	Streptococcus muta
8	73	64.6	1475 23 AAU98027	S. mutans glucosyl
9	73	64.6	1475 23 AAU98030	S. mutans glucosyl

10	73	64.6	1475 23 AAU98031	S. mutans glucosyl
11	73	64.6	1475 23 AAU98032	S. mutans glucosyl
12	73	64.6	1475 23 AAU98033	S. mutans glucosyl
13	73	64.6	1475 23 AAU98034	S. mutans glucosyl
14	73	64.6	1475 23 AAU98035	S. mutans glucosyl
15	73	64.6	1475 23 AAU98036	S. mutans glucosyl
16	73	64.6	1475 23 AAU98037	S. mutans glucosyl
17	73	64.6	1475 23 AAU98038	S. mutans glucosyl
18	73	64.6	1475 23 AAU98039	S. mutans glucosyl
19	73	64.6	1475 23 AAU98040	S. mutans glucosyl
20	73	64.6	1475 23 AAU98041	S. mutans glucosyl
21	73	64.6	1577 17 AAU91047	Alpha-D-glucosyltr
22	71	62.8	1375 23 AAU98028	S. mutans glucosyl
23	71	62.8	1375 23 AAU79288	Streptococcus muta
24	66	58.4	12 23 ABB98643	L. mesenteroides a
25	65	57.5	2057 21 AAB10667	Dextran saccharase
26	64	56.6	2835 23 ABB98574	Glucosyltransferase
27	62	54.9	1592 14 AAR32925	Dextran saccharase
28	59	52.2	12 23 ABB98641	Dextran-saccharas
29	58	51.3	1527 23 AAU80055	Dextran-saccharas
30	53	46.9	12 23 ABB98642	Dextran-saccharas
31	52	46.0	12 23 ABB98646	Dextran-saccharas
32	52	46.0	12 23 ABB98647	Dextran-saccharas
33	51	45.1	12 23 ABB98647	Dextran-saccharas
34	50	44.2	121 22 AAU60582	Protonibacterium
35	49	43.4	1429 23 ABB47797	Yeast DNA topoisom
36	48.5	42.9	218 21 AAB43277	Human ORFX ORF3041
37	48.5	42.9	289 22 AAM25871	Human protein sequ
38	48.5	42.9	319 22 AAU32156	Novel human secret
39	48.5	42.9	427 20 AAU75761	Human secreted pro
40	48.5	42.9	427 21 AAB32412	Human secreted pro
41	48.5	42.9	436 21 AAB32411	Human secreted pro
42	48.5	42.9	576 22 AAB32384	Human protein sequ
43	48.5	42.9	576 22 AAB94297	Human protein sequ
44	48.5	42.9	576 22 ABB94569	Human protein sequ
45	48.5	42.9	12 23 ABB98578	Dextran saccharas

ALIGNMENTS

RESULT 1
AAU98029 standard; Protein; 1430 AA.
AC AAU98029;
XX 27-AUG-2002 (first entry)
XX S. mutans glucosyltransferase GTFP.
XX DE Glucosyltransferase; GTFP; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amylolast; vacuole; paper manufacture.
XX KW Streptococcus mutans.
XX OS US2002031826-A1.
XX PN 14-MAR-2002.
XX PD 19-DEC-2000; 2000US-0740274.
XX PF 11-DEC-1998; 98US-0210361.
XX PR 07-JUN-1995; 95US-0478704.
XX PR 07-JUN-1995; 95US-0482711.
XX PR 07-JUN-1995; 95US-0485243.
XX PR 16-JAN-1998; 98US-0007999.
XX PR 16-JAN-1998; 98US-0008172.
XX PR 20-JAN-1998; 98US-0009620.
XX PA (NICH/) NICHOLS S E.

PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR N-PSDB; ABR52940.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Disclosure; Page 38-42; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild-type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch.
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFD.
 CC
 CC
 CC Sequence 1430 AA;
 XX
 XX
 XX Query Match 66.4%; Score 75; DB 23; Length 1430;
 XX Best Local Similarity 75.0%; Pred. No. 0.0077; 4; Indels 0; Gaps 0;
 XX Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 XX
 XX 1 AIDHSLTEAWSGNDNDYVK 20
 XX ||:|||||||
 XX Db 495 AINHSLTEAWSGNDNDYVK 514
 XX
 XX
 XX RESULT 2
 XX AAU98041
 XX ID AAU98041 standard; Protein; 1430 AA.
 XX AC AAU98041;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFD mutant T589D.
 XX
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.
 XX

OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX FT MISC-difference 589 /note="Wild-type Thr substituted by Asp"
 XX FT US2002031826-A1.
 XX
 XX PD 14-MAR-2002.
 XX
 XX PF 19-DEC-2000; 2000US-0740274.
 XX
 XX PR 11-DEC-1998; 98US-0210361.
 XX PR 07-JUN-1995; 95US-0478704.
 XX PR 07-JUN-1995; 95US-0482711.
 XX PR 07-JUN-1995; 95US-0485243.
 XX PR 16-JAN-1998; 98US-0007999.
 XX PR 16-JAN-1998; 98US-0008172.
 XX PR 20-JAN-1998; 98US-0009620.
 XX
 XX PA (NICH)/ NICHOLS S E.
 XX
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild-type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch.
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC Note: The present sequence represents a GTF mutant of the invention.
 CC and the information in claim 36.
 CC

SQ Sequence 1430 AA;
 Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AIDHLSILEAMSGNDNDYK 20
 DB 495 AINHLSTLEAMSDNDPQYK 514
 RESULT 3
 AAU98042 ID AAU98042 standard; Protein; 1430 AA.
 AC AAU98042;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant T589E.
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 589
 FT /note= "Wild-type Thr substituted by Glu"
 FT US2002031826-A1.
 EN 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI: 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT Claim 36; Page -; 44pp; English.
 PS The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y110A/Y117A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.
 SQ Sequence 1430 AA;
 Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AIDHLSILEAMSGNDNDYK 20
 DB 495 AINHLSTLEAMSDNDPQYK 514
 RESULT 4
 AAU98043 ID AAU98043 standard; Protein; 1430 AA.
 XX AAU98043;
 AC AAU98043;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant N471D.
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 471
 FT /note= "Wild-type Asn substituted by Asp"
 FT US2002031826-A1.
 EN 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI: 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT Claim 36; Page -; 44pp; English.
 PS The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y110A/Y117A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFD sequence appearing as AAU98029
CC and the information in claim 36.

XX Sequence 1430 AA;

Query Match 66.4%; Score 75; DB 23; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.007;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
||:|||||
DB 495 AINHLISLEAMSDNDPQYK 514

RESULT 6
AAU98045
ID AAU98045 standard; Protein; 1430 AA.

XX AAU98045;

XX 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFD mutant N471D/T589E.

XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
XX coating composition; glucan; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutlein.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 471 /note= "Wild-type Asn substituted by Asp"

XX Misc-difference 589 /note= "Wild-type Thr substituted by Glu"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substitutes for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
XX (GTF) B polypeptide having changes at position from I448V, D457N,
XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799G/K1014T,
XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
XX an isolated polynucleotide which encodes P1 or P2, or its complementary
XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTF mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFD sequence appearing as AAU98029
CC and the information in claim 36.

XX Sequence 1430 AA;

Query Match 66.4%; Score 75; DB 23; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.007;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
||:|||||
DB 495 AINHLISLEAMSDNDPQYK 514

RESULT 7
AAU79285
ID AAU79285 standard; Protein; 1017 AA.

XX AAU79285;

XX 13-AUG-2002 (first entry)

XX Streptococcus mutans monoclonal antibody-related protein #2.

XX Antibody; dental caries; water insoluble glucan synthetase;

XX anti-carries; glucosyl transferase-B; immunotherapy.

XX Streptococcus mutans.

XX JP2002114709-A.

XX 16-APR-2002.

XX 04-OCT-2000; 2000JP-0304889.

XX 04-OCT-2000; 2000JP-0304889.

XX (UYNI-) UNIV NIPPON.

XX WPI; 2002-448101/48.

XX Anti-carries agent composed of a monoclonal antibody against an
XX inhibitory enzyme against water insoluble glucan synthetase of glucosyl
XX transferase-B (GTF-B) of Streptococcus mutans -

PS Claim 4; Page 17-19; 28pp; Japanese.

XX The invention relates to a monoclonal antibody against dental caries and

CC an anti-carries agent composed of a monoclonal antibody produced by

CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)

CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having

CC inhibitive activity against water insoluble glucan synthetase of glucosyl

CC transferase-B. The monoclonal antibody specifically inhibits water

CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl

CC transferase-B and is used in the immunotherapy of dental caries. This

CC sequence represents a Streptococcus mutans monoclonal antibody-related

CC protein.

XX

XX Sequence 1017 AA;

SO

Query Match 64.6%; Score 73; DB 23; Length 1017;

Best Local Similarity 73.7%; Pred. No. 0.0098;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYV 19

DB 447 ANDHLSILEAWSGNDNDYV 465

RESULT 8

AAU98027

ID AAU98027 standard; Protein; 1475 AA.

XX

AC AAU98027;

XX

DT 27-AUG-2002 (first entry)

XX

DE S. mutans glucosyltransferase GTFB.

XX

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture.

XX

OS Streptococcus mutans.

XX

PN US2002031826-A1.

XX

PD 14-MAR-2002.

XX

PF 19-DEC-2000; 2000US-0740274.

XX

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX

PA (NICH/) NICHOLS S E.

XX

PI Nichols SE;

XX

DR WP1; 2002-414332/44.

DR N-FSDB; ABR52938.

XX

PT Glucosyltransferase B or D protein useful for producing a glucan useful

PT as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions -

XX

PS Disclosure; Page 21-25; 44pp; English.

XX

XX The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from I448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,

CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having

CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the

CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

CC coating composition comprising a glucan produced in a plant transformed

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where

CC the glucan is produced in the amyloplast and/or vacuole or a maize line

CC deficient in starch biosynthesis, transformed with a gene encoding a

CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful

CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing

CC conditions to produce a regenerated plant and inducing expression of the

CC polynucleotide for a time sufficient to produce the glucan in the

CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC is produced in the amyloplast of potato or the vacuole of sugar beet.

CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper

CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC properties and impart gloss to the paper during coating step.

CC The present sequence represents GTFB.

XX

SO Sequence 1475 AA;

Query Match 64.6%; Score 73; DB 23; Length 1475;

Best Local Similarity 73.7%; Pred. No. 0.015;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYV 19

DB 481 ANDHLSILEAWSGNDNDYV 499

RESULT 9

AAU98030

ID AAU98030 standard; Protein; 1475 AA.

XX

AC AAU98030;

XX

DT 27-AUG-2002 (first entry)

XX

DE S. mutans glucosyltransferase GTFB mutant I448V.

XX

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; muten.

XX

OS Streptococcus mutans.

XX

PN Synthetic.

XX

FT Key

FT Msc-difference 448

PN US2002031826-A1.

XX

PD 14-MAR-2002.

XX

PF 19-DEC-2000; 2000US-0740274.

XX

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 PS
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 64.6%; Score 73; DB 23; Length 1475;
 Best Local Similarity 73.7%; Pred. NO. 0.015; Indels 0; Gaps 0;
 Matches 14; Conservative 1; Mismatches 4;
 QY 1 AIDHSLTEAMSGNDNDYV 19
 Db 481 ANDHSLTEAMSGNDNDYVYL 499
 RESULT 10
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX

DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457
 PT /note= "Wild-type Asp substituted by Asn"
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PE 19-DEC-2000; 2000US-0740274.
 XX
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 PS
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

XX
XX Sequence 1475 AA;
SQ

Query Match 64.6%; Score 73; DB 23; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSLEAWSGNDNDYV 19
DB 481 ANDHLSLEAWSNDNDPYL 499

RESULT 11
AAU98032 ID AAU98032 standard; Protein; 1475 AA.
XX
XX AAU98032;
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX S. mutans glucosyltransferase GTFB mutant D567T.
DE
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutein.
KM
XX
XX Streptococcus mutans.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 567 /note="Wild-type Asp substituted by Thr"
FT
XX
XX US2002031826-A1.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 19-DEC-2000; 2000US-0740274.
PF
XX
XX 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
PA
XX
XX Nichols SE;
PI
XX
XX WPI; 2002-414332/44.
DR
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44p; English.
PS
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, KI014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/KI014T, 1448V/D457N/D567T/D571K/KI014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

SQ Sequence 1475 AA;

Query Match 64.6%; Score 73; DB 23; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSLEAWSGNDNDYV 19
DB 481 ANDHLSLEAWSNDNDPYL 499

RESULT 12
AAU98033 ID AAU98033 standard; Protein; 1475 AA.
XX
XX AAU98033;
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX S. mutans glucosyltransferase GTFB mutant KI014T.
DE
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutein.
KM
XX
XX Streptococcus mutans.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 1014 /note="Wild-type Iys substituted by Thr"
FT
XX
XX US2002031826-A1.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 19-DEC-2000; 2000US-0740274.
PF
XX
XX 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/ NICHOLS S E.
 XX Nichols SE,
 XX WPI; 2002-414332/44.
 DR
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in Claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 64.6%; Score 73; DB 23; Length 1475;
 Best Local Similarity 73.7%; Pred. No. 0.015;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

AC AAU98034;
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N/D567T.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX amyloplast; vacuole; paper manufacture; mutant; mutan.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 XX MISC-difference 457
 XX MISC-difference 567
 XX /note= "Wild-type Asp substituted by Asn"
 XX MISC-difference 567
 XX /note= "Wild-type Asp substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/ NICHOLS S E.
 XX Nichols SE,
 XX WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
 XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX an isolated polynucleotide which encodes P1 or P2, or its complementary
 XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX an expression cassette comprising the polynucleotide operably linked to a
 XX promoter, a vector comprising the expression cassette, host cell
 XX introduced with the vector, a transgenic plant comprising the
 XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 XX coating composition comprising a glucan produced in a plant transformed
 XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
 XX thermoplastic molecule or their combinations or glucan and starch where
 XX the glucan is produced in the amyloplast and/or vacuole or a maize line
 XX deficient in starch biosynthesis, transformed with a gene encoding a
 XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 XX comprising the glucan (paper sizing/coating agent). The vector is useful
 XX for producing a glucan in a plant. The method comprises transforming a
 XX plant cell with the vector, growing the plant cell under plant growing
 XX conditions to produce a regenerated plant and inducing expression of the
 XX polynucleotide for a time sufficient to produce the glucan in the
 XX regenerated plant, where the vector contains a transit sequence from
 XX ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 XX chlorophyll AB binding protein to produce a transgenic plant, and glucan
 XX is produced in the amyloplast of potato or the vacuole of sugar beet.

US2002031826-A1.
14-MAR-2002.
19-DEC-2000; 2000US-0740274.
11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.
07-JUN-1995; 95US-0485243.
16-JAN-1998; 98US-0007999.
16-JAN-1998; 98US-0008172.
20-JAN-1998; 98US-0009620.
(NICH) NICHOLS S E.
Nichols SE;
WPI; 2002-414332/44.
Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
Claim 36; Page -; 44pp; English.
The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T, Y169A/Y170A/Y171A, and K799Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AA098027 and the information in claim 36.

Sequence 1475 AA;

Query Match 64.6%; Score 73; DB 23; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

0Y 1 AIDHLSILEAWSGNDNDYV 19

DB 481 AIDHLSILEAWSGNDNDYV 499

Search completed: November 13, 2003, 09:38:27
Job time : 38.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds

(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-13

Perfect score: 113

Sequence: 1 AIDHSLTEAMSGNDNDYKQ 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	95.6	1365	2	A41483
2	75	66.4	1431	2	A45866
3	73	64.6	1475	2	B31135
4	73	64.6	1577	2	T30858
5	71	62.8	1375	2	JT0345
6	69	61.1	1508	2	T31098
7	69	61.1	1518	2	A44811
8	67	59.3	1449	2	T30857
9	67	59.3	1449	2	T30552
10	62	54.9	1592	2	A38175
11	61	54.0	1599	2	S22737
12	49	43.4	2392	2	B84176
13	49	43.4	1428	1	IS8712
14	48	42.5	366	2	A83211
15	48	42.5	378	2	T04254
16	48	42.5	378	2	B72129
17	48	42.5	695	2	F66493
18	48	42.5	704	2	C81542
19	48	42.5	3206	1	GN58PV
20	47	41.6	298	2	T15906
21	47	41.6	680	2	T39858
22	47	41.6	769	2	D86335
23	46	40.7	287	2	T36413
24	46	40.7	458	2	H86314
25	46	40.7	514	2	F87592
26	46	40.7	536	2	T13261
27	46	40.7	1252	2	H97178
28	46	40.7	1802	2	T00020
29	45	39.8	2366	2	S10317

30	45	39.8	2367	2	S70172	toxin B - Clostrid
31	44	38.9	342	2	E69581	acetoin dehydrogen
32	44	38.9	389	2	T17601	hypothetical prote
33	44	38.9	468	2	S61964	probable membrane
34	44	38.9	588	2	C83836	subtilisin-type pr
35	43.5	38.5	356	2	T48354	hypothetical prote
36	43	38.1	76	2	T42309	hypothetical prote
37	43	38.1	162	2	C97754	hypothetical prote
38	43	38.1	177	2	T40825	hypothetical prote
39	43	38.1	221	2	A87411	DNA-binding respon
40	43	38.1	270	2	AC2987	conserved hypothet
41	43	38.1	270	2	E98296	hypothetical prote
42	43	38.1	272	2	H72596	hypothetical prote
43	43	38.1	424	2	AD2238	hypothetical prote
44	43	38.1	429	2	T45040	hypothetical prote
45	43	38.1	697	2	H71525	probable outer mem

ALIGNMENTS

RESULT 1

A41483
Glucosyltransferase (EC 2.4.1.1) gtfS precursor - Streptococcus sobrinus
C.Species: Streptococcus sobrinus
C.Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C.Accession: A41483
R.Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
Infect. Immun. 58, 2452-2458, 1990
A.Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase
A.Reference number: A41483; MUID:90316665; PMID:2142479
A.Accession: A41483
A.Molecule type: DNA
A.Residues: 1-1365 <GIL>
A.Cross-references: GB:M0943; NID:G153652; PIDN:AAA26898.1; PID:G153653
C.Genetics:
A.Gene: gtfS
C.Superfamily: cpl repeat homology
C.Keywords: glycosyltransferase; hexosyltransferase

Query Match 95.6%; Score 108; DB 2; Length 1365;
Best local similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDHSLTEAMSGNDNDYK 20
DB 467 AIDHSLTEAMSGNDNDYK 486

RESULT 2

A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C.Species: Streptococcus mutans
C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C.Accession: A45866
R.Honda, O.; Kato, C.; Kuramitsu, H.K.
J. Gen. Microbiol. 136, 2099-2105, 1990
A.Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase
A.Reference number: A45866; MUID:9110958; PMID:2146600
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1431 <HON>
A.Cross-references: GB:M29296
C.Superfamily: cpl repeat homology
C.Keywords: glycosyltransferase; hexosyltransferase
F.181-201/Domain: cpl repeat homology <CP1>
F.1127-1146/Domain: cpl repeat homology <CP2>
F.1192-1211/Domain: cpl repeat homology <CP3>
F.1257-1276/Domain: cpl repeat homology <CP4>
F.1277-1297/Domain: cpl repeat homology <CP5>
F.1321-1340/Domain: cpl repeat homology <CP6>
F.1341-1361/Domain: cpl repeat homology <CP6>

F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 66.4%; Score 75; DB 2; Length 1431;
 Best Local Similarity 75.0%; Pred. No. 0.0023;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
 ||:|||||:|||||:|
 DB 495 AINHLSTLEAMSDNDPQYK 514

RESULT 3

gtfB protein precursor - Streptococcus mutans
 C:Species: Streptococcus mutans
 C>Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999

C:Accession: B33135; A33128
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987

A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685

A:Accession: B33135
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-1475 <SH1>

A:Cross-references: GB:M17361; NID:G153639; PIDN:AAA88588.1; PID:G153640
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.

Submitted to the Protein Sequence Database, September 1990
 A:Reference number: A33128

A:Accession: A33128
 A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA
 A:Residues: 1-171,173-641,'N',643-1475 <SH2>

A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology

F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>

F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>

F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 64.6%; Score 73; DB 2; Length 1475;
 Best Local Similarity 73.7%; Pred. No. 0.0049;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 19
 ||:|||||:|||||:|
 DB 481 ANDHLSILEAMSGNDTPYL 499

RESULT 4

glucosyltransferase - Streptococcus salivarius
 C:Species: Streptococcus salivarius
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
 R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
 Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for p1
 A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-1577 <SIM>

A:Cross-references: EMBL:L55928; NID:G662380; PID:G662381; PIDN:AA04143.1
 C:Genetics: gtfm

A:Gene: gtfm

Query Match 64.6%; Score 73; DB 2; Length 1577;
 Best Local Similarity 75.0%; Pred. No. 0.0053;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20

DB 588 AIAHLSILEAMSYNDHQYK 607
 ||:|||||:|||||:|

RESULT 5

UT0345
 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)

N:Alternate names: sucrose 6-glucosyltransferase
 C:Species: Streptococcus mutans

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999

C:Accession: UT0345; C33135
 R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.

Gene 69, 101-109, 1988

A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
 A:Reference number: UT0345; MUID:89137980; PMID:2976010

A:Accession: UT0345
 A:Molecule type: DNA

A:Residues: 1-1375 <UED>
 A:Experimental source: GS-5

R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987

A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685

A:Accession: C33135
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-349 <SH1>

A:Cross-references: GB:M17361
 C:Genetics: gtfC

C:Function: catalyzes the synthesis of both water-soluble and water-insoluble glucans

C:Superfamily: cpl repeat homology
 C:Keywords: duplication; glycosyltransferase; hexosyltransferase

F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-137/Product: glucosyltransferase #status predicted <MAT>

F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1252-1272/Domain: cpl repeat homology <CP2>

F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 62.8%; Score 71; DB 2; Length 1375;
 Best Local Similarity 73.7%; Pred. No. 0.0094;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 19
 ||:|||||:|||||:|
 DB 507 ANDHLSILEAMSYNDTPYL 525

RESULT 6
 probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T31098
 R:Nonchols, V.; Renaud-Simeon, M.; Monan, P.; Willemot, R.M.
 FEBS Microbiol. Lett. 159, 307-315, 1998

A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A:Reference number: Z20981; MUID:98164374; PMID:9503626

A:Accession: T31098
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-1508 <MON>

A:Cross-references: EMBL:AF030129; NID:G2766611; PID:G2766612; PIDN:AA095453.1
 A:Experimental source: strain NRRL B-1299

C:Genetics: dsrB
 A:Gene: dsrB

A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C:Keywords: glycosyltransferase, hexosyltransferase

Query Match 61.1%; Score 69; DB 2; Length 1508;
 Best Local Similarity 70.0%; Pred. No. 0.022;

QY 1 AIDHLSILEAMSGNDNDYK 20

Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSIEAWSGNDNDYK 20
| | | | |
| | | | |
Db 563 ANQHLSIEAWSHNDPEYK 582

RESULT 7

A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
C.Species: Streptococcus salivarius
C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1995
C.Accession: A44811; S22726; S28809
R.Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A.Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
A.Reference number: A44811; PMID:92148377; PMID:1838391
A.Accession: A44811
A.Molecule type: DNA
A.Residues: 1-1518 <GIF>
A.Cross-references: EMBL:Z11873; NID:G47526; PIDN:CAA77900.1; PID:G47527
A.Note: sequence extracted from NCBI backbone (NCBIN:81050, NCBIIP:81052)
C.Genetics:
A.Gene: gtfI
C.Superfamily: cpl repeat homology
C.Keywords: glycosyltransferase; hexosyltransferase
F.1307-1326/Domain: cpl repeat homology <CP4>

Query Match 61.1%; Score 69; DB 2; Length 1518;
Best Local Similarity 66.7%; Pred. No. 0.022;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 AIDHLSIEAWSGNDNDY 18
| | | | |
| | | | |
Db 501 ALAHSIVAEWSNDNH 518

RESULT 8
T30857
glucosyltransferase - Streptococcus salivarius
C.Species: Streptococcus salivarius
C.Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C.Accession: T30857
R.Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A.Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pili
A.Reference number: Z20909; PMID:95122197; PMID:7822030
A.Accession: T30857
A.Molecule type: DNA
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-1449 <SIM>
A.Cross-references: EMBL:L35495; NID:G662378; PID:G662379; PIDN:AAC41412.1
C.Genetics:
A.Gene: gtfI

Query Match 59.3%; Score 67; DB 2; Length 1449;
Best Local Similarity 77.8%; Pred. No. 0.043;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDHLSIEAWSGNDNDY 18
| | | | |
| | | | |
Db 536 AIKHSIEAWSHNDAY 553

RESULT 9
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C.Species: Streptococcus salivarius
C.Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C.Accession: T30552
R.Jaffe, R.I.
submitted to the EMBL Data Library, February 1998
A.Description: Streptococcus salivarius V1477 gtfN.

A.Reference number: Z20854
A.Accession: T30552
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-1449 <JAF>
A.Cross-references: EMBL:AF049609; NID:G2933545; PID:G2933546; PIDN:AAC05156.1
C.Genetics:
A.Gene: gtfN

Query Match 59.3%; Score 67; DB 2; Length 1449;
Best Local Similarity 77.8%; Pred. No. 0.043;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDHLSIEAWSGNDNDY 18
| | | | |
| | | | |
Db 536 AIKHSIEAWSHNDAY 553

RESULT 10
A8175
glucosyltransferase precursor - Streptococcus sobrinus
C.Species: Streptococcus sobrinus
C.Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
C.Accession: A8175
R.Abo, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
J. Bacteriol. 173, 989-996, 1991
A.Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A.Reference number: A8175; PMID:91123227; PMID:1704006
A.Accession: A8175
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1592 <ABO>
A.Cross-references: GB:D90213; NID:G217033; PIDN:BA14241.1; PID:G1014946; PID:G217033
C.Superfamily: cpl repeat homology
F.1093-1112/Domain: cpl repeat homology <CP1>
F.1122-1241/Domain: cpl repeat homology <CP2>
F.1287-1306/Domain: cpl repeat homology <CP3>
F.1330-1351/Domain: cpl repeat homology <CP4>
F.1352-1371/Domain: cpl repeat homology <CP5>
F.1402-1420/Domain: cpl repeat homology <CP6>
F.1465-1484/Domain: cpl repeat homology <CP7>
F.1513-1532/Domain: cpl repeat homology <CP8>

Query Match 54.9%; Score 62; DB 2; Length 1592;
Best Local Similarity 57.9%; Pred. No. 0.3;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDHLSIEAWSGNDNDY 19
| | | | |
| | | | |
Db 477 ANHVSIVAEWSNDTPYL 495

RESULT 11
S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
C.Species: Streptococcus salivarius
C.Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C.Accession: S22737; S28810; B44811; S22727
R.Jacques, N.
submitted to the EMBL Data Library, March 1992
A.Reference number: S22737
A.Accession: S22737
A.Molecule type: DNA
A.Residues: 1-1599 <JAC>
A.Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAA77898.1; PID:G47531
A.Experimental source: ATCC 25975
R.Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A.Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
A.Reference number: A44811; PMID:92148377; PMID:1838391
A.Accession: S28810
A.Molecule type: DNA
A.Residues: 1-51 <GIF>

A:Cross-references: EMBL:Z11873
 C:Genetics:
 A:Gene: glfK
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 P:1456-1475/domain: cpl repeat homology <CPR>

Query Match 54.0%; Score 61; DB 2; Length 1599;
 Best Local Similarity 57.1%; Pred. No. 0.43;
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AIDHLSLEAMSGNDNDYVQ 21
 DB 491 ALANISLEAMSHNDPYVNE 511

RESULT 12

Hypothetical protein Vng0156c [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: B84176
 R:Ng, W.V.; Kennedy, S.P.; Mahlars, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.; Jaitnauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jaidic, Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: B84176
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <STO>
 A:Cross-references: GB:AE004437; NID:910579804; PIDN:AAG18774.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: VNG0156C

Query Match 43.4%; Score 49; DB 2; Length 232;
 Best Local Similarity 47.4%; Pred. No. 3.8;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 AIDHLSLEAMSGNDNDYV 19
 DB 62 AADPVRSYDAMSGRDADHV 80

RESULT 13

ISBYT2
 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) - yeast (Saccharomyces cerevisiae)
 N:Alternate names: DNA gyrase; DNA topoisomerase II; protein N2244; protein YNL088w
 C:Species: Saccharomyces cerevisiae
 C:Date: 30-Sep-1992 #sequence_revision 10-Nov-1995 #text_change 16-Jun-2000
 C:Accession: S57534; A25630; S63027; S30866; S65093
 R:Solier-Mira, A.; Saliz, J.E.; Ballesta, J.P.G.; Remacha, M.
 submitted to the EMBL Data Library, June 1995
 A:Reference number: S57533
 A:Accession: S57534
 A:Molecule type: DNA
 A:Residues: 1-1428 <SOL>
 A:Cross-references: EMBL:X89016; NID:9887621; PIDN:CAA61422.1; PID:9887623
 R:Glaever, G.; Lynn, R.; Goto, T.; Wang, J.C.
 U:Biol. Chem. 261, 12448-12454, 1986
 A:Title: The complete nucleotide sequence of the structural gene TOP2 of yeast DNA topoi
 A:Reference number: A25630; MUID:86304413; PMID:3017975
 A:Accession: A25630
 A:Molecule type: DNA
 A:Residues: 1-74, 'N', 75-546, 'L', 548-836, 'R', 838-1428 <GIA>
 A:Cross-references: GB:M13814; NID:9172997; PIDN:AA36610.1; PID:9172998
 R:Solier-Mira, A.; Saliz, J.E.; Ballesta, J.P.G.; Remacha, M.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63018
 A:Accession: S63018
 A:Molecule type: DNA

A:Residues: 1-1428 <SOM>
 A:Cross-references: EMBL:Z71364; NID:91301988; PIDN:CAA95964.1; PID:91301989; GSPDB:GN00
 A:Experimental source: strain S288C
 R:Pannatipour, M.; Liu, Y.X.; Nittes, J.L.
 submitted to the EMBL Data Library, January 1993
 A:Description: The top2-5 mutant of yeast topoisomerase II encodes an enzyme resistant t
 A:Reference number: S30866

A:Accession: S30866
 A:Molecule type: DNA
 A:Residues: 812-836, 'R', 838-882, 'P', 884, 'T', 887-977 <JAN>
 A:Cross-references: EMBL:L08968; NID:9172999; PIDN:AA59328.1; PID:9173000
 R:Solier-Mira, A.; Saliz, J.E.; Ballesta, J.P.G.; Remacha, M.
 Yeast 12, 485-491, 1996
 A:Title: The sequence of a 17 933 bp segment of Saccharomyces cerevisiae chromosome XIV
 A:Reference number: S65092; MUID:96310628; PMID:8740422

A:Accession: S65093
 A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Cross-references: EMBL:X89016; NID:9887621; PIDN:CAA61422.1; PID:9887623
 C:Comment: Type II DNA topoisomerase catalyzes the ATP-dependent transient breakage, pas

C:Genetics: SGD:TOP2; TOR3; TRF3; TOP2-5; MIPS:YNL088w

A:Cross-references: SGD:S0005032; MIPS:YNL088w

A:Map position: 14L

C:Superfamily: eukaryotic type II DNA topoisomerase; phase T4 DNA topoisomerase (ATP-hy
 C:Keywords: ATP; DNA binding; DNA replication; heterotetramer; isomerase; phosphoprotein

Query Match 43.4%; Score 49; DB 1; Length 1428;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 4 HSLSLAMSGNDNDYV 19
 DB 623 HMKIFSLQGNDRXYI 638

RESULT 14

AE3211
 Hypothetical protein potF [imported] - Agrobacterium tumefaciens (strain C58, Dupont) p1
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AE3211
 R:Wood, D.W.; Serubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Gunthner, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Tung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, B.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AE3211
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-366 <KIR>
 A:Cross-references: GB:AE008687; PIDN:AA146107.1; PID:917743872; GSPDB:GN00188
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: potF
 A:Genome: plasmid

Query Match 42.5%; Score 48; DB 2; Length 366;
 Best Local Similarity 52.9%; Pred. No. 9.2;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 DHSLSLAMSGNDNDYV 19
 DB 179 DFLSLMLYQGNDDYV 195

RESULT 15

TO4254
 Hypothetical protein F20B18.100 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C:Accession: T04254
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hohnesiel, J.; Newes, H.W.; Mayer, K.F.X.
Submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15263
A:Accession: T04254
A:Molecule type: DNA
A:Residues: 1-378

A:Cross-references: EMBL:AL049483
A:Experimental source: cultivar Columbia; BAC clone F20B18
C:Genetics:
A:Map position: 4
A:Introns: 326/3
A:Note: F20B18.100

Query Match	42.5%;	Score 48;	DB 2;	Length 378;
Best Local Similarity	41.2%;	Pred. No. 9.6;		
Matches	7;	Conservative	6;	Mismatches 4;
				Indels 0;
				Gaps 0;

```
QY      2 IDHLSILEAWSGNDNDY 18
      ::: ||||| :::
Db     295 LDYDGVLEAWSGKESPF 311
```

```
Search completed: November 13, 2003, 09:50:27
Job time : 14.0379 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049A-13

Sequence: 1 AIDHLSILEAWSGNDNDYKQ 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

Database : SwissProt_41: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Length	DB	ID	Description
1	108	95.6	1365	1	GTF5_STPDO	P29336 streptococc	
2	75	66.4	1462	1	GTFD_STRMU	P49331 streptococc	
3	73	64.6	1476	1	GTFB_STRMU	P08387 streptococc	
4	71	62.8	1455	1	GTFC_STRMU	P13470 streptococc	
5	62	54.9	1592	1	GTF2_STPDO	P27470 streptococc	
6	62	54.9	1597	1	GTF1_STPDO	P11001 streptococc	
7	49	43.4	1428	1	TOP2_YEAST	P06786 saccharomye	
8	48	42.5	3206	1	POLG_PSMAN	P29152 p genome po	
9	48	42.5	6885	1	SNE2_HUMAN	P08w10 homo sapien	
10	47	41.6	680	1	YH2X_SCHPO	O7443 schizosacch	
11	45	39.8	2366	1	TOXB_CLODI	P18177 clostridium	
12	44	38.9	341	1	ACOB_BASUS	O34591 bacillus au	
13	44	38.9	487	1	STK4_HUMAN	O13043 homo sapien	
14	43	38.1	572	1	PEMU_DICDI	O23919 dictyostell	
15	43	38.1	674	1	COA4_BACTU	O08905 bacillus th	
16	43	38.1	766	1	SC15_DROME	O9vade6 drosophila	
17	43	38.1	1161	1	RPO2_FOPPV	O91544 fowipox viri	
18	42.5	37.2	272	1	Y028_BUCAI	P57140 buchnera ap	
19	42	37.2	794	1	Y966_CONGL	P35667 corynebacte	
20	41.5	36.7	128	1	Y532_BUCAI	P57588 buchnera ap	
21	41.5	36.7	151	1	Y347_METJA	O57793 methanococc	
22	41.5	36.7	680	1	NCPB_CANTR	P37201 candida tro	
23	41.5	36.7	844	1	PHSG_DROME	O9xrt19 drosophila	
24	41	36.3	116	1	Y960_HA8IN	P44084 haemophilus	
25	41	36.3	295	1	VENB_VIBVU	B74965 vibrio vlin	
26	41	36.3	396	1	PEK_LEPIN	B8f58b lepidospira	
27	41	36.3	666	1	POL_FMYD	P09523 f1gwort mos	
28	41	36.3	801	1	PIPA_DICDI	O02158 dictyostell	
29	41	36.3	808	1	PKH1_AKNSP	O8yw11 anabaena sp	
30	41	36.3	923	1	K685_MOUSE	O8w332 mus musculu	
31	41	36.3	927	1	K685_HUMAN	O75170 homo sapien	
32	41	36.3	1141	1	UBP2_SCHPO	O9p300 schizosacch	
33	41	36.3	1634	1	POL2_METJA	O58295 methanococc	

34	40.5	35.8	310	1	P1R1_ARATH	G91P6 arabidopsis
35	40.5	35.8	833	1	CW1_YEAST	Q3108 saccaromyc
36	40.5	35.8	1461	1	TOP2_CANAL	P87078 candida alb
37	40	35.4	53	1	YORU_TV1	P19305 thermoprote
38	40	35.4	121	1	RUS_SOME	P93779 solanum mel
39	40	35.4	179	1	FLJA_SALAB	P52618 salmoneilla
40	40	35.4	179	1	FLJA_SALAB	P52619 salmoneilla
41	40	35.4	179	1	FLJA_SALTY	P52619 salmoneilla
42	40	35.4	261	1	UPK_XYIFA	Q9PCE0 xyliella fas
43	40	35.4	261	1	UPK_XANCP	O6PD65 xanthomonas
44	40	35.4	265	1	UPK_XANNC	O6PD66 xanthomonas
45	40	35.4	302	1	CH14_SOLTU	P52406 solanum tub

ALIGNMENTS

RESULT 1	
GTFS_STRODO	
ID GTFS_STRODO	STANDARD; PRT; 1365 AA.
AC	P29336;
DT	01-DEC-1992 (Rel. 24, Created)
DT	01-DEC-1992 (Rel. 24, last sequence update)
DT	28-FEB-2003 (Rel. 41, last annotation update)
DE	glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE	(Sucrose 6-glucosyltransferase).
GN	GIFS.
OS	Streptococcus downei (Streptococcus sobrinus).
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC	Streptococcus.
CC	NCBL_TaxID=1317;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	STRAIN=MPE28;
RA	MEDLINE=9031665; PubMed=2142479;
RL	Glimore K.S., Russell R.R., Ferretti J.J.;
RT	"Analysis of the Streptococcus downei gtfS gene, which specifies a
RT	glucosyltransferase that synthesizes soluble glucans.";
RL	Infect. Immun. 58:2452-2458 (1990).
CC	- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC	- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC	fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC	- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC	PRIMER GLUCAN UNLIKE GTF-I.
CC	- DISEASE: DENTAL CARIES.
CC	- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC	1,6-GLUCOSE).
CC	- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUTAN-
CC	- BINDING PROTEIN FROM S.MUTANS.
CC	- SIMILARITY: Contains 10 cell wall binding repeats.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb.ch/announce/
CC	or send an email to license@isb.ch).
CC	-----
DR	EMBL; M30943; AAA26898.1; -
DR	InterPro; IPR003479; CW binding.
DR	InterPro; IPR003479; Glyco_hydro_70.
DR	Pfam; PF01473; CW binding 1; 8.
DR	Pfam; PF0324; Glyco_hydro_70; 1.
KW	transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT	SIGNAL 1
FT	CHAIN 36
FT	REPEAT 37 1365
FT	REPEAT 157 177
FT	REPEAT 178 197
FT	DOMAIN 198 1061
FT	CATALYTIC (APPROXIMATE).

FT REPEAT 1062 1082 CELL WALL BINDING 3.
 FT REPEAT 1083 1102 CELL WALL BINDING 4.
 FT REPEAT 1150 1169 CELL WALL BINDING 5.
 FT REPEAT 1170 1190 CELL WALL BINDING 6.
 FT REPEAT 1225 1243 CELL WALL BINDING 7.
 FT REPEAT 1289 1308 CELL WALL BINDING 8.
 FT REPEAT 1309 1328 CELL WALL BINDING 9.
 FT REPEAT 1331 1352 CELL WALL BINDING 10.
 SQ SEQUENCE 1365 AA; 151590 MW; 1672965A2B8C476 CXC64;

Query Match 95.6%; Score 108; DB 1; Length 1365;
 Best Local Similarity 100.0%; Pred. No. 6.4e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDHSLTEAWSGNDNDYK 20
 DB 467 AIDHSLTEAWSGNDNDYK 466

RESULT 2
 GTFD_STRMU STANDARD; PRT; 1462 AA.
 ID GTFD_STRMU STANDARD; PRT; 1462 AA.
 AC PA9331; 069383; 069386; 069389; 069392; 069398;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DN (Sucrose 6-glucosyltransferase).
 GN GTFD OR SMU.910.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RA MEDLINE=91100958; PubMed=2148600;
 RA Honda O., Kato C., Kuramitsu H.K.;
 RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
 the glucosyltransferase-S enzyme";
 RL J. Gen. Microbiol. 136:2099-2105(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 RC MT4467 / Serotype E, and MT8148 / Serotype C;
 RA MEDLINE=98231643; PubMed=9570124;
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of
 Streptococcus mutans";
 RL FEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / Serotype C;
 RA MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Iain S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl](N) = D-
 CC fructose + [(1,6)-alpha-D-glucosyl](N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M29296; AAA26895.1; -
 CC EMBL; D88653; BAA26103.1; -
 CC EMBL; D88656; BAA26107.1; -
 CC EMBL; D88659; BAA26111.1; -
 CC EMBL; D88662; BAA26115.1; -
 CC EMBL; D89979; BAA26121.1; -
 CC EMBL; AE014932; AAN58619.1; -
 CC InterPro; IPR002479; CW binding.
 CC InterPro; IPR003318; Glyco_hydro_70.
 CC Pfam; PF02324; Glyco_hydro_70; I
 CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
 CC Complete proteome.
 CC SIGNAL 1 1462
 CC CHAIN ? 1462
 CC DOMAIN 1232 1423
 CC REPEAT 1232 1295
 CC REPEAT 1296 1359
 CC REPEAT 1360 1423
 CC VARIANT 10 10
 CC VARIANT 19 19
 CC VARIANT 58 58
 CC VARIANT 68 68
 CC VARIANT 81 81
 CC VARIANT 113 113
 CC VARIANT 122 122
 CC VARIANT 132 132
 CC VARIANT 135 135
 CC VARIANT 137 137
 CC VARIANT 202 202
 CC VARIANT 255 255
 CC VARIANT 275 275
 CC VARIANT 288 288
 CC VARIANT 301 301
 CC VARIANT 313 313
 CC VARIANT 317 317
 CC VARIANT 328 328
 CC VARIANT 350 350
 CC VARIANT 628 633
 CC VARIANT 688 688
 CC VARIANT 726 732
 CC VARIANT 726 730
 CC VARIANT 762 762
 CC VARIANT 964 964
 CC VARIANT 1019 1019
 CC VARIANT 1059 1060
 CC VARIANT 1060 1060
 CC VARIANT 1080 1080
 CC VARIANT 1142 1142
 CC VARIANT 1198 1198
 CC VARIANT 1220 1220
 CC VARIANT 1280 1280
 CC VARIANT 1282 1282
 CC -----
 CC POTENTIAL.
 CC GLUCOSYLTRANSFERASE-S.
 CC 3 X 63 AA APPROXIMATE TANDEM REPEATS.
 CC 1.
 CC 2.
 CC 3.
 CC Y -> H (IN STRAINS GS-5, MT4239, MT4245,
 CC MT4251, MT4467 AND MT8148).
 CC I -> V (IN STRAINS GS-5, MT4239, MT4245,
 CC MT4251, MT4467 AND MT8148).
 CC K -> E (IN STRAIN MT4467).
 CC A -> S (IN STRAINS MT4239 AND MT4245).
 CC A -> T (IN STRAINS MT4251 AND MT8148).
 CC T -> I (IN STRAINS MT4239 AND MT4245).
 CC A -> V (IN STRAINS MT4239, MT4245 AND
 CC MT8148).
 CC A -> S (IN STRAINS GS-5 AND MT4467).
 CC A -> V (IN STRAIN MT4245).
 CC A -> T (IN STRAINS GS-5, MT4239, MT4245,
 CC MT4251, MT4467 AND MT8148).
 CC V -> L (IN STRAIN MT4239).
 CC D -> N (IN STRAIN MT8148).
 CC E -> D (IN STRAINS MT4239, MT4245 AND
 CC MT4251).
 CC D -> N (IN STRAINS MT4239, MT4245 AND
 CC MT4251).
 CC Q -> H (IN STRAIN MT4245).
 CC D -> N (IN STRAINS MT4239 AND MT4251).
 CC E -> K (IN STRAIN MT4239).
 CC V -> F (IN STRAIN MT4239).
 CC F -> L (IN STRAINS MT4239, MT4251 AND
 CC MT4467).
 CC KKYYTQ -> EKEYTL (IN STRAIN MT4251).
 CC A -> S (IN STRAIN MT4239).
 CC TDGSEA -> ADKNGDS (IN STRAIN MT4251).
 CC TDGGS -> ADKKN (IN STRAINS MT4239 AND
 CC MT4245).
 CC T -> A (IN STRAINS GS-5, MT4239, MT4245,
 CC MT4251, MT4467 AND MT8148).
 CC D -> Y (IN STRAIN MT4251).
 CC E -> K (IN STRAINS MT4245 AND MT4251).
 CC LG -> IR (IN STRAIN MT4251).
 CC G -> R (IN STRAIN MT4245).
 CC G -> Q (IN STRAIN MT4239).
 CC H -> R (IN STRAIN GS-5).
 CC S -> C (IN STRAIN MT4239).
 CC Y -> C (IN STRAINS MT4251 AND MT4467).
 CC F -> L (IN STRAIN MT4467).
 CC Q -> P (IN STRAIN MT4245).

-I- SIMILARITY: Contains 10 cell wall binding repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce> or send an email to license@isb.sib.ch).

EMBL; M17361; AAA8858.1; -
EMBL; D88651; BAA26101.1; -
EMBL; D88654; BAA26105.1; -
EMBL; D88657; BAA26109.1; -
EMBL; D88660; BAA26113.1; -
EMBL; D89777; BAA26119.1; -
EMBL; AE014940; AAN58705.1; -
InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco hydro_70.
PFam; PF001473; CW binding_1; 1.
DR Pfam; PF02324; Glyco hydro_70; 1.
KM Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
Complete proteome.

SIGNAL
1 34
FT CHAIN 35 1476
FT DOMAIN 35 1051
FT DOMAIN 1097 1476
FT REPEAT 1097 1130
FT DOMAIN 1161 1470
FT REPEAT 1161 1210
FT REPEAT 1225 1275
FT REPEAT 1290 1340
FT REPEAT 1355 1405
FT REPEAT 1420 1470
FT REPEAT 62 62
FT VARIANT 65 65
FT VARIANT 68 68
FT VARIANT 78 78
FT VARIANT 86 86
FT VARIANT 89 89
FT VARIANT 168 168
FT VARIANT 276 276
FT VARIANT 399 399
FT VARIANT 474 474
FT VARIANT 512 512
FT VARIANT 519 519
FT VARIANT 701 701
FT VARIANT 708 708
FT VARIANT 938 938
FT VARIANT 952 957
FT VARIANT 963 964
FT VARIANT 968 970
FT VARIANT 1086 1086
FT VARIANT 1158 1158
FT VARIANT 1163 1163
FT VARIANT 1168 1168
FT VARIANT 1182 1182
FT VARIANT 1234 1234
FT VARIANT 1263 1263
FT VARIANT 1264 1264
FT VARIANT 1272 1272
FT VARIANT 1329 1329
FT VARIANT 1394 1394

POTENTIAL.
GLUCOSYLTRANSFERASE-I.
CATALYTIC (APPROXIMATE).
GLUCAN-BINDING (APPROXIMATE).
A REPEAT.
5 X TANDEM REPEATS.
1.
2.
3.
4.
5.
S -> T (IN STRAIN MT4239).
T -> I (IN STRAIN GS-5).
V -> A (IN STRAINS GS-5, MT4245, MT4251,
MT4467 AND MT8148).
Q -> P (IN STRAIN MT4251).
I -> S (IN STRAINS GS-5, MT4245, MT4251,
MT4467 AND MT8148).
S -> F (IN STRAIN MT4251).
K -> N (IN STRAIN MT4251).
S -> D (IN STRAINS GS-5, MT4467 AND
MT8148).
N -> R (IN STRAIN MT4239).
I -> T (IN STRAIN MT4239).
K -> R (IN STRAIN MT8148).
F -> Y (IN STRAIN MT8148).
T -> I (IN STRAIN MT8148).
A -> V (IN STRAIN MT8148).
F -> L (IN STRAIN MT8148).
FGPEVE -> VGPEVA (IN STRAINS GS-5, MT4239
AND MT4467).
SV -> NT (IN STRAINS GS-5, MT4239 AND
MT4467).
ADS -> VDQG (IN STRAINS GS-5, MT4239 AND
MT4467).
A -> T (IN STRAIN MT4239).
S -> N (IN STRAIN MT4239).
H -> Y (IN STRAIN MT4251).
E -> K (IN STRAIN MT8148).
Y -> C (IN STRAIN MT8148).
A -> P (IN STRAIN MT4239).
R -> H (IN STRAIN MT8148).
R -> H (IN STRAINS GS-5 AND MT4467).
Y -> H (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
S -> G (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
H -> Y (IN STRAINS GS-5 AND MT4467).
Y -> H (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).

```
FT  VARIANT 1402 1402 S -> G (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT 1459 1459 AND MT8148).
FT  CONFLICT 570 570 Y -> H (IN STRAIN MT4467).
FT  CONFLICT 800 817 AODVAVASTASTDGG -> LKMFALRLARPHQWA
FT  CONFLICT (IN REF. 1).
FT  CONFLICT 1310 1310 H -> L (IN REF. 1).
SQ  SEQUENCE 1476 AA, 165846 MW, 9C6E09F731B4C6CF CRC64;

Query Match 64.6%; Score 73; DB 1; Length 1476;
Best Local Similarity 73.7%; Pred. No. 0.0024;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHSLTEAMSGNDNDY 19
Db 481 ANDHSLTEAMSGNDNDY 499

RESULT 4
GTRC_STRMU STANDARD; PRT; 1455 AA.
AC P13470; 069382; 069385; 069388; 069397; P05427;
AT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN GTF OR SWU.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RV [1]
RV SEQUENCE FROM N.A.
RP STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 691101-109(1988).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT4148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujisawa T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Nejar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RN SEQUENCE OF 1-349 FROM N.A.
RP STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
```



```
FT VARIANT 1377 1377 R -> K (IN STRAIN MT8148).
FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT CONFLICT 1337 1455 ORLYKSGVQAKGKHLITERGRKRYDPENSGNRYRYR
TSSGMYTFRGNDGALIGMHVREGRRVYFDENGVRYASHD
QRNMDYDRDFGRSSAVRFRSRNGFDFNFRF ->
HASILSLMVFRLRSSLSQSVKVSNTMLLPEMKFVLM
(IN REF. 1).
SQ SEQUENCE 1455 AA; 162965 MW; 3CB455A99A4FEC86 CRC64;

Query Match 62.8%; Score 71; DB 1; Length 1455;
Best Local Similarity 73.7%; Pred. No. 0.0048;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 AIDHSLTEAWSGNDNDYV 19
Db 507 ANDHSLTEAWSYNDTPYL 525

RESULT 5
GTF2 STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase)
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BAA14241.1; -
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CW_binding_I; 13.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC
```

```
KW transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7.
SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351ECF CRC64;

Query Match 54.9%; Score 62; DB 1; Length 1592;
Best Local Similarity 57.9%; Pred. No. 0.14;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 AIDHSLTEAWSGNDNDYV 19
Db 477 ANNHSLTEAWSYNDTPYL 495

RESULT 6
GTF1 STRDO STANDARD; PRT; 1597 AA.
AC P11001;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.F., Gilpin M.L., Russell R.R.B.;
RA "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RA sobrinus MFe28.";
RL J. Bacteriol. 169:4271-4278(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M7391; AAC63063.1; -
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC
```

```

DR Pfam: PF01473; CW binding_1, 16.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38
FT CHAIN 39 1597
FT DOMAIN 39 1050
FT DOMAIN 1099 1597
FT DOMAIN 1099 1597
FT REPEAT 1099 1132
FT REPEAT 1163 1213
FT REPEAT 1227 1277
FT REPEAT 1282 1342
FT REPEAT 1352 1399
FT REPEAT 1406 1455
FT REPEAT 1465 1512
FT REPEAT 1519 1568
FT REPEAT 1582 1597
SQ SEQUENCE 1597 AA; 177080 MW; B9B6A200868798E CRC64;

Query Match
Best Local Similarity 54.9%; Score 62; DB 1; Length 1597;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHSLIEFAMSGNDIV 19
Db 483 ANNHVSIIVEAWSNDTPYL 501

RESULT 7
TOP2_YEAST STANDARD; PRT; 1428 AA.
AC P06786;
DT 01-JAN-1988 (Rel. 06; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
GN TOP2 OR TOR3 OR YNL088W OR N2244.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86304413; PubMed=3017975;
RA Glaeser F., Lynn R., Goto T., Wang J.C.;
RT "The complete nucleotide sequence of the structural gene TOP2 of
RT Yeast DNA topoisomerase II."
RL J. Biol. Chem. 261:12448-12454(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RX MEDLINE=96310628; PubMed=8740422;
RA Soler-Mira A., Sait J.E., Ballestra J.P.G., Remacha M.;
RT "The sequence of a 17,933 bp segment of Saccharomyces cerevisiae
RT chromosome XIV contains the RHO2, TOP2, MKT1 and END3 genes and five
RT new open reading frames."
RL Yeast 12:485-491(1996).
RN [3]
RP REVIEW ON PHOSPHORYLATION.
RX MEDLINE=93073815; PubMed=1332607;
RX Glaeser S.M., Walter R., Dang O., Cardenas M.E.;
RA Berger J.M., Gambin S.J., Harrison S.C., Wang J.C.;
RT "Structure and mechanism of DNA topoisomerase II."
RL Nature 379:225-232(1996).
-1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC MAKES DOUBLE-STRAND BREAKS. ESSENTIAL DURING MITOSIS AND MEIOSIS
CC FOR PROPER SEGREGATION OF DAUGHTER CHROMOSOMES.

CC CC
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: PHOSPHORYLATION ENHANCES THE ACTIVITY. IT STIMULATES ITS
CC DECATENATION ACTIVITY.
CC -1- MISCELLANEOUS: IN YEAST TOPOISOMERASE II CAN SUBSTITUTE
CC TOPOISOMERASE I FOR THE RELAXING ACTIVITY.
CC -1- MISCELLANEOUS: Eukaryotic topoisomerase I and II can relax both
CC negative and positive supercoils, whereas prokaryotic enzymes
CC relax only negative supercoils.
CC -1- SIMILARITY: Belongs to the type II topoisomerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: M13814; AB36610.1; -
CC EMBL: X89016; CA61422.1; -
CC EMBL: Z71364; CA95964.1; -
CC PIR: S57534; ISBYT2.
CC PDB: 1BGW; 1I-JUL-96.
CC PDB: 1BUT; 04-MAY-99.
CC SGD: S005032; TOP2.
CC
CC GO: GO:0005634; C:nucleus, IDA.
CC GO: GO:0005716; C:synaptonemal complex, IDA.
CC GO: GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity, IDA.
CC GO: GO:0006333; P:chromatin assembly/disassembly, IMP.
CC GO: GO:0006271; P:DNA strand elongation; IMP.
CC GO: GO:0006265; P:DNA topological change, IDA.
CC GO: GO:0007131; P:meiotic recombination, IMP.
CC GO: GO:000020; P:negative regulation of recombination within. . .; IMP.
CC InterPro: IPR003594; ATPbind_ATPase.
CC InterPro: IPR003957; CBFA_NFYB_top1s.
CC InterPro: IPR001241; DNA_topoisolt.
CC InterPro: IPR002205; DNA_topoisolt.
CC
CC Pfam: PF00204; DNA_gyraseb; 1.
CC Pfam: PF00521; DNA_topoisolt; 1.
CC Pfam: PF02518; HATPase_c; 1.
CC PRINTS: PR00615; COAATSUBUNITA.
CC PRINTS: PR00418; TP12FAMILY.
CC ProDom: PD000742; DNA_topoisolt; 1.
CC SMART: SM00387; HATPase_c; 1.
CC SMART: SM00433; TOP2c; 1.
CC SMART: SM00434; TOP4c; 1.
CC PROSITE: PS00177; TOPOISOMERASE II; 1.
CC KX Isomerase; Topoisomerase; DNA-binding; ATP-binding; Phosphorylation;
CC Nuclear protein; 3D-structure.
CC NP_BIND 140 145
CC ACT_SITE 782 782
CC MOD_RES 1086 1086
CC MOD_RES 1087 1087
CC MOD_RES 1258 1258
CC MOD_RES 1266 1266
CC MOD_RES 1269 1269
CC MOD_RES 1272 1272
CC MOD_RES 1353 1353
CC MOD_RES 1356 1356
CC MOD_RES 1408 1408
CC MOD_RES 1423 1423
CC CONFLICT 74 74
CC CONFLICT 547 547
CC CONFLICT 837 837
CC TURN 427 428
CC STRAND 430 430
CC TURN 433 436
CC TURN 438 439
CC HELIX 440 442
CC STRAND 444 445

```

```

FT    HELIX    450  464
FT    STRAND   469  474
FT    HELIX    489  502
FT    TURN    503  505
FT    TURN    509  510
FT    TURN    513  514
FT    STRAND   520  524
FT    HELIX    533  545
FT    HELIX    547  550
FT    TURN    551  551
FT    TURN    553  554
FT    STRAND   556  558
FT    STRAND   564  569
FT    STRAND   573  578
FT    HELIX    581  590
FT    TURN    591  591
FT    HELIX    592  594
FT    STRAND   597  601
FT    HELIX    610  627
FT    TURN    628  629
FT    HELIX    633  691
FT    STRAND   694  694
FT    TURN    695  697
FT    HELIX    701  712
FT    TURN    713  713
FT    STRAND   719  720
FT    HELIX    721  732
FT    TURN    733  733
FT    HELIX    738  749
FT    TURN    753  754
FT    TURN    760  761
FT    STRAND   764  764
FT    HELIX    770  772
FT    TURN    773  773
FT    TURN    775  776
FT    TURN    780  782
FT    STRAND   784  785
FT    TURN    789  790
FT    TURN    791  794
FT    HELIX    797  802
FT    STRAND   805  810
FT    TURN    809  815
FT    STRAND   815  824
FT    HELIX    824  833
FT    STRAND   838  841
FT    STRAND   844  844
FT    HELIX    846  858
FT    TURN    859  859
FT    STRAND   867  867
FT    TURN    870  871
FT    STRAND   875  878
FT    TURN    881  882
FT    STRAND   883  887
FT    STRAND   889  894
FT    TURN    895  896
FT    STRAND   897  902
FT    TURN    905  906
FT    HELIX    909  920
FT    STRAND   932  935
FT    STRAND   943  946
FT    HELIX    949  958
FT    HELIX    960  963
FT    TURN    964  965
FT    STRAND   967  971
FT    TURN    973  974
FT    STRAND   975  978
FT    TURN    980  981
FT    STRAND   984  987
FT    HELIX    990  1034
FT    TURN    1035  1036
FT    HELIX    1045  1054
FT    TURN    1055  1056

```

```

FT    STRAND   1059  1060
FT    TURN    1062  1063
FT    STRAND   1066  1067
FT    TURN    1108  1110
FT    STRAND   1112  1112
FT    HELIX    1114  1117
FT    TURN    1118  1118
FT    HELIX    1121  1124
FT    HELIX    1126  1147
FT    TURN    1148  1148
FT    HELIX    1151  1175

```

Query Match 43.4%; Score 49; DB 1; Length 1428;
 Best Local Similarity 50.0%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HSLTEAMSGNDNDYV 19
 Db 623 HLKIFSLQNDKDYI 638

RESULT 8
 POLG_PSBMV STANDARD; PRT; 3206 AA.
 AC P29152;

DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polypeptide [Contains: N-terminal protein (p1); Helper
 component proteinase (EC 3.4.22.45) (HC-Pro); Protein p3; 6 kDa
 protein 1 (6k1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
 (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
 (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
 inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
 (EC 2.7.7.48); Coat protein (CP)].
 DE Pea seed-borne mosaic virus (strain DPDI).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
 OC Polyvirus.
 NCBI_TaxID=31736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9204431; PubMed=1940858;
 RA Johansen E., Rasmussen O.F., Heide M., Borkhardt B.;
 RT "The complete nucleotide sequence of pea seed-borne mosaic virus
 RNA.";
 RL J. Gen. Virol. 72:2625-2632 (1991).
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
 further restricted by preferences for the amino acids in p6 - p1,
 that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
 Glu+(Ser or Gly) for the enzyme from tobacco etch virus. The
 natural substrate is the viral polypeptide, but other proteins and
 oligopeptides containing the appropriate consensus sequence are
 also cleaved.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 {RNA} (N).
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|Gly bond at its own C-
 terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|Gly, in the
 processing of the polyviral polypeptide.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -1- PTM: THE VIRAL RNA OF POLYVIRUSES IS EXPRESSED AS A SINGLE
 POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL, PROTEOLYTIC
 PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POLYVIRUSES POLYPEPTIDE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).

DR EMBL: D10930; BAA01726.1; -
 DR PIR: J01331; GNVSFV.
 DR MEROPS: C04.010; -
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR001730; Peptidase_C4.
 DR InterPro: IPR001456; Peptidase_C6.
 DR InterPro: IPR001592; Poly_coat.
 DR InterPro: IPR002540; Poly_P1.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3.
 DR InterPro: IPR007094; RNA_pol_PSV1r.
 DR Pfam: PF00271; Helicase_C; 1.
 DR Pfam: PF00863; Peptidase_C4; 1.
 DR Pfam: PF00851; Peptidase_C6; 1.
 DR Pfam: PF00767; Poly_coat; 1.
 DR Pfam: PF01577; Poly_P1; 1.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR PRINTS: PR00966; NIAPO1PTASE.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR HydroLase; Transferase; 1.
 KM Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
 KM ATP-binding.
 FT CHAIN 1 ? N-TERMINAL PROTEIN.
 FT CHAIN ? 856 HELPER COMPONENT PROTEINASE.
 FT CHAIN ? ? PROTEIN P3.
 FT CHAIN ? 1266 6 kDa PROTEIN 1.
 FT CHAIN 1267 1902 CYTOPLASMIC INCLUSION PROTEIN.
 FT CHAIN 1903 1955 6 kDa PROTEIN 2.
 FT CHAIN 1956 ? GENOME-LINKED PROTEIN.
 FT CHAIN ? 2395 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 2396 2915 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 2916 3206 COAT PROTEIN.
 FT CHAIN 1266 1267 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 1902 1903 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 1955 1956 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 2395 2396 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 2915 2916 CLEAVAGE (BY 49 kDa PROTEASE).
 FT BINDING 2016 COVALENT LINKAGE OF VIRAL RNA (BY
 FT SIMILARITY).
 FT NP BIND 1351 1358 ATP (POTENTIAL).
 SO SEQUENCE 3206 AA; 364271 MW; 42A3D921BB940CEP CRC64;
 QY 2 IDHLS--LEAMSGNDNDVYK 20
 DB 394 IDHFSIVGIKIWNAPDAEYK 414
 Query Match 42.5%; Score 48; DB 1; Length 3206;
 Best Local Similarity 47.6%; Pred. No. 51;
 Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND
 RP INTERACTION WITH F-ACTIN.
 RX MEDLINE=22113122; PubMed=12118075;
 RA Zhen Y.-Y., Libotte T., Munck M., Noegel A.A., Korenbaum E.;
 RT "NUANCE, a giant protein connecting the nucleus and actin
 RT cytoskeleton." J. Cell Sci. 115:3207-3222(2002).
 RL J. Cell Sci. 115:3207-3222(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 4; 5 AND 7).
 RX MEDLINE=21652858; PubMed=11792814;
 RA Zhang Q., Skepper J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G.,
 RT Weisberg P.L., Ellis J.A., Shanahan C.M.;
 RT "Nesprins: a novel family of spectrin-repeat-containing proteins that
 RT localize to the nuclear membrane in multiple tissues." J.
 RL J. Cell Sci. 114:4485-4498(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22296983; PubMed=12408964;
 RA Zhang Q., Kagnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;
 RT "The nesprins are giant actin-binding proteins, orthologous to
 RT Drosophila melanogaster muscle protein MSP-300." J.
 RL Genomics 80:473-481(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 6).
 RX TISSUE-Testis.
 RA Pousetka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL [5]
 RP SEQUENCE FROM N.A.
 RX PubMed=12508121;
 RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
 RA Catolico L., Levy M., Barbe V., de Bernardis V., Ureca-vial A.,
 RA Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,
 RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Crnaud C.,
 RA Brulets T., Jallion O., Friedlander L., Samson G., Broctier P.,
 RA Cure S., Segures B., Aniere F., Saman S., Crespeau H., Abbasi N.,
 RA Alich N., Bocus D., Dickhoff R., Dors M., Dubois I., Friedman C.,
 RA Gouyenoux M., James R., Madan A., Mairey-Batista B., Mangelot S.,
 RA Martins N., Menard M., Oztas S., Ratcliffe A., Shafer T., Trask B.,
 RA Vacherie B., Bellemere C., Belser C., Besnard-Gonnet M.,
 RA Barol-Mavel D., Boulard M., Briet-Silla S., Combette S.,
 RA Dufosse-Laurent V., Ferron C., Lechaplats C., Louesse C., Museliet D.,
 RA Magdelanet G., Patteu F., Petit E., Stryain-Trukiewicz P., Trybou A.,
 RA Vega-Czarny N., Batallie E., Bluet E., Bordelais I., Dubois M.,
 RA Dumont C., Guerin T., Hafray S., Hammadi R., Munga J., Pellouin V.,
 RA Robert D., Wunderle E., Gauguier G., Roy A., Sainte-Marthe L.,
 RA Verdier J., Verdier-Discala C., Hillier L., Fulton L., Moperson J.,
 RA Matsuda F., Wilson R., Scarpetti C., Gyapay G., Winkler P., Saurin W.,
 RA Quetier F., Waterston R., Hood L., Weissbach J.;
 RT "The DNA sequence and analysis of human chromosome 14." J.
 RL Nature 421:601-607(2003).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 8).
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Soares K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitsugu S., Cantini P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.C., Malek A.M., Gnatatue P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield A.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [7]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6885
 RP FROM N.A.
 RC TISSUE=Splice, and Tongue;
 RA Jikuya H., Takano U., Nomura N., Kikuno R., Nagase T., Ohara O.,
 RA Nimomiyu K., Magatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Futuya T., Takahashi M., Kikawa E., Omura Y., Abe K., Kamihara K.,
 RA Kasuya N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Kawai-Hio Y., Sato K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
 RT "NDO human cDNA sequencing project";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RP [8]
 RP SEQUENCE OF 5709-6885 FROM N.A. (ISOFORM 2).
 RP TISSUE=Brain;
 RC MEDLINE=99246063; PubMed=10231033;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroseawa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro";
 RL DNA Res. 6:63-70(1999).
 RP [9]
 RP REVISIONS.
 RC TISSUE=Brain;
 RX MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual
 RT curation of 330 KIAA cDNA clones";
 RL DNA Res. 9:99-106(2002).
 RP [10]
 RP SEQUENCE OF 5754-6885 FROM N.A.
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gaassenhuber J., Glaesl S.,
 RA Ansoyge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber U., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.-W., Oltenevaider B., Obermaier B., Tampe J., Heubner D.,
 RA Mambert R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs";
 RL Genome Res. 11:422-435(2001).
 CC -1- FUNCTION: Involved in the maintenance of nuclear organization and
 CC structural integrity. Probably anchoring protein which tethers the
 CC nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton
 CC by interacting with the nuclear envelope and with F-actin in the
 CC cytoplasm.
 CC -1- SUBUNIT: Interacts with F-actin via its N-terminal domain.
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential). The
 CC largest part of the protein is cytoplasmic, while its C-terminal
 CC part is associated with the nuclear envelope, most probably the
 CC outer nuclear membrane. Remains associated with the nuclear
 CC envelope during its breakdown in mitotic cells.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=9;
 CC Name=1;
 CC IsoId=Q8WXH0-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8WXH0-2; Sequence=VSP_007164, VSP_007166;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q8WXH0-3; Sequence=VSP_007155;
 CC Note=Produced by exon skipping that results in a frameshift. No
 CC experimental confirmation available;

CC Name=4; Synonyms=Beta;
 CC IsoId=Q8WXH0-4; Sequence=VSP_007156;
 CC Name=5; Synonyms=Alpha;
 CC IsoId=Q8WXH0-5; Sequence=VSP_007157, VSP_007164, VSP_007165;
 CC Name=6;
 CC IsoId=Q8WXH0-6; Sequence=VSP_007158, VSP_007165, VSP_007166;
 CC Note=No experimental confirmation available;
 CC Name=7; Synonyms=Gamma;
 CC IsoId=Q8WXH0-7; Sequence=VSP_007154, VSP_007163;
 CC Name=8;
 CC IsoId=Q8WXH0-8; Sequence=VSP_007161, VSP_007162;
 CC Note=No experimental confirmation available;
 CC Name=9; Synonyms=NUANCE-N-33;
 CC IsoId=Q8WXH0-9; Sequence=VSP_007159, VSP_007160;
 CC -1- TISSUE SPECIFICITY: Widely expressed, with higher level in kidney,
 CC adult and fetal liver, stomach and placenta. Weakly expressed in
 CC skeletal muscle and brain. Isoform 5 is highly expressed in
 CC pancreas, skeletal muscle and heart.
 CC -1- DOMAIN: The Klarsicht domain mediates the nuclear envelope
 CC targeting.
 CC -1- SIMILARITY: Belongs to the Nesprin family.
 CC -1- SIMILARITY: Contains 1 actin-binding domain.
 CC -1- SIMILARITY: Contains 2 calponin-homology (CH) domains.
 CC -1- SIMILARITY: Contains 1 Klarsicht domain.
 CC -1- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC -1- SIMILARITY: Contains 9 spectrin repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 CC EMBL: AF435010; AAL33547.1; -
 CC DR EMBL: AF435011; AAL33548.1; -
 CC DR EMBL: AY061757; AAL33800.1; -
 CC DR EMBL: AY061758; AAL33801.1; -
 CC DR EMBL: AY061759; AAL33802.1; -
 CC DR EMBL: AF435911; AAN60443.1; -
 CC DR EMBL: AL117404; CAB55905.1; -
 CC DR EMBL: AL117404; CAB55905.1; -
 CC DR EMBL: AL152832; -; NOT_ANNOTATED_CDS.
 CC DR EMBL: AL355094; -; NOT_ANNOTATED_CDS.
 CC
 CC Query Match 42.5%; Score 48; DB 1; Length 6885;
 CC Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 CC Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 3 DHSIHEAWSGNDN 16
 CC Db 362 DHLQREAWDGDH 375
 CC
 CC RESULT 10
 CC YH2X SCHPO STANDARD; PRT; 680 AA.
 CC ID YH2X SCHPO
 CC AC 074343;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Hypothetical UPF0024 protein CIA4.09 in chromosome II.
 CC GN SPBC1A4.09.
 CC OS Schizosaccharomyces pombe (fission yeast).
 CC CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC OC NCBI_TaxID=4896;
 CC OX 11;
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21848401; PubMed=11859360;
 CC Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagals K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
 RA Money P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton G., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
 RA Welfens I., Vanstreets E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelarge V., Mottier S.,
 RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Paga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong U., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shipkovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe."
 RA Nature 415:871-880(2002).
 CC -1- SIMILARITY: Belongs to the UPF0024 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL031174; CAA20114.1; -
 DR PIR; T39858; T39858.
 DR GeneDB; Spombe; SPBC1A.09; -
 DR InterPro; IPR001656; UPF0024.
 DR Pfam; PF01142; UPF0024; 1.
 DR TIGRPFAMs; TIGR00094; TIGR00094; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 680 AA; 76438 MW; B4A52DAECBA1B9A6 CRC64;
 Query Match 41.6%; Score 47; DB 1; Length 680;
 Best Local Similarity 62.5%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
 QY 6 SILEAM--SGNDNDYV 19
 DB 434 SILEWTSRSGNCTDYL 449
 RESULT 11
 TOXB CLODI
 ID TOXB CLODI STANDARD; PRT; 2366 AA.
 AC P18177;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Toxin B.
 GN TOXB OR TCDB.
 OS Clostridium difficile.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI 10463;
 RX MEDLINE=90326540; PubMed=2374729;
 RA Barroto L.A., Wang S.Z., Phelps C.J., Johnson U.L., Wilkins T.D.,
 RT "Nucleotide sequence of Clostridium difficile toxin B gene."
 RL Nucleic Acids Res. 18:4004-4004(1990).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI 10463;
 RA von Eichel-Streiber C.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1271-2366 FROM N.A.
 RC STRAIN=VPI 10463;
 RX MEDLINE=92293124; PubMed=1603068;
 RA Eichel-Streiber C., Laufenberg-Feldmann R., Sarlingen S., Schlze J.,
 RA Sauerborn M.,
 RT "Comparative sequence analysis of the Clostridium difficile toxins A
 RT and B."
 RL Mol. Gen. Genet. 233:260-268(1992).
 CC -1- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN ENTEROTOXIN
 CC CALLED A AND CYTOTOXIN B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X53138; CAA37298.1; -
 DR EMBL; X92982; CAA63562.1; -
 DR EMBL; X60984; CAA43299.1; -
 DR PIR; A27636; A27636.
 DR PIR; S10317; S10317.
 DR InterPro; IPR002479; CW binding.
 DR Pfam; PF01473; CW binding 1; 17.
 DR Pfam; PF04488; Gly_transf_sug; 1.
 KW Toxin.
 SQ SEQUENCE 2366 AA; 269709 MW; E1024BDB8A56ADF CRC64;
 Query Match 39.8%; Score 45; DB 1; Length 2366;
 Best Local Similarity 44.4%; Pred. No. 1;le=02;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 2 IDHSILEAMSGNDNDYV 19
 DB 1733 INDLSIRYWSNDGNDPI 1750
 RESULT 12
 ACOB BACSU
 ID ACOB BACSU STANDARD; PRT; 341 AA.
 AC O34591;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acetoin:2,6-dichlorophenolindophenol oxidoreductase beta subunit
 DE (EC 1.1.1.-) (Acetoin:DCPIP oxidoreductase-beta) (AO:DCPIP OR)
 DE (TTP-dependent acetoin dehydrogenase E1 beta-subunit).
 GN ACOB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=9296597; PubMed=10368162;
 RA Huang M., Oppermann-Sanio F.B., Steibuchel A.,
 RT "Biochemical and molecular characterization of the Bacillus subtilis
 RT acetoin catabolic pathway."
 RL J. Bacteriol. 181:3837-3841(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AC327;
 RX MEDLINE=97124190; PubMed=8969503;
 RA Yamamoto H., Uchiyama S., Sekiguchi J.,
 RT "Cloning and sequencing of a 40.6 kb segment in the 73 degrees-76

RT degrees region of the Bacillus subtilis chromosome containing genes
 RT for trehalose metabolism and acetoin utilization.";
 RL Microbiology 142:3057-3065(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brignelli S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conterton I.F., Cummings N.J., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Erington J., Fabret C., Ferrati E., Fougere D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grand G.,
 RA Giuseppe G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Ilaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa K., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Priesen E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha F., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takuchi M., Yamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: CATALYZES THE 2, 6-DICHLOROPHENOLINDOPHENOL-DEPENDENT
 CC CLEAVAGE OF ACETONIN INTO ACETATE AND ACETALDEHYDE (BY SIMILARITY).
 CC -1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: TO THE BETA SUBUNITS OF 2-OXO-ACID DEHYDROGENASE
 CC COMPONENTS OF VARIOUS MULTISUBUNIT COMPLEXES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF006075; AAC05583.1; -
 CC EMBL; D78509; BAA24295.1; -
 CC EMBL; Z99108; CAB12636.1; -
 CC PIR; B69581; B69581.
 CC HSSP; P09061; 1080.
 CC Subtilist; B612559; acob.
 CC InterPro; IPR005476; Transketolase_C.
 CC InterPro; IPR005475; Transketolase_CR.
 CC Pfam; PF02779; transket_pyr; 1.
 CC Pfam; PF02780; transketolase_C; 1.
 CC Acetoin catabolism; Oxidoreductase; Complete proteome.
 CC INT MET 0
 CC BY SIMILARITY.
 CC SEQUENCE 341 AA; 36713 MM; DCSB3D549C89EFO CRC64;
 CC
 CC Query Match 38.9%; Score 44; DB 1; Length 341;
 CC Best Local Similarity 61.5%; Pred. No. 17;
 CC Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 36 AYDHLQDDEAWG 48
 ||| |||
 RESULT 13
 ID STK4_HUMAN STANDARD; PRT; 487 AA.
 AC Q13043; Q15802; Q9NT24;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STK20-like kinase
 DE MST1) (MST-1) (Mammalian STK20-like protein kinase 1)
 DE (Serine/threonine protein kinase Krs-2).
 OS STK4 Or MST1.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394929; PubMed=7665586;
 RA Creasy C.L., Chernoff J.;
 RT "Cloning and characterization of a human protein kinase with homology
 RT to Ste20".
 RL J. Biol. Chem. 270:21695-21700(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413604; PubMed=8816758;
 RA Taylor L.K., Wang H.C., Erikson R.L.;
 RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
 RT 2".
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
 RN [3]
 RP SEQUENCE OF 1-435 FROM N.A.
 RL Laird G.;
 CC Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
 CC -1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
 CC THE KINASE ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- PM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SRE20 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U18297; AAA83254.1; -
 CC EMBL; U60207; AAB17262.1; -
 CC EMBL; AL109839; CAB89421.1; -
 CC HSSP; P24941; IHCL.
 CC Genew; HGNC:11408; STK4.
 CC MIM; 604965; -
 CC GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC InterPro; IPR001245; Tyr_kinase.
 CC Pfam; PF00069; pkinase; 1.
 CC PRINTS; PR00109; TYRKINASE.
 CC PRODOM; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

Transferrase; Serine/threonine-protein kinase; ATP-binding.
 KM DOMAIN 30 281 PROTEIN KINASE.
 FT NP BIND 36 44 ATP (BY SIMILARITY).
 FT BINDING 59 59 ATP (BY SIMILARITY).
 FT ACT_SITE 149 149 BY SIMILARITY.
 FT DOMAIN 373 378 POLY-GLU.
 FT CONFLICT 222 222 P -> R (IN REF. 1).
 FT CONFLICT 312 312 V -> M (IN REF. 1).
 SO SEQUENCE 487 AA; 55630 MW; 150758B5C5F77D5C CRC64;

Query Match 38.9%; Score 44; DB 1; Length 487;
 Best Local Similarity 61.5%; Pred. No. 26;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 EAMSGNDNDYKQ 21
 Db 248 ELMSDNDYKQ 260

RESULT 14
 PGMT_DICDI
 ID PGMT_DICDI STANDARD; PRT; 572 AA.
 AC Q23919;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphoglucosyltransferase (EC 5.4.2.2) (Glucose phosphotransferase) (PGM).
 GN PGM.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Loomis W.F.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS ENZYME PARTICIPATES IN BOTH THE BREAKDOWN AND SYNTHESIS OF GLUCOSE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose 1-phosphate = alpha-D-glucose 6-phosphate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOHEXOSE MUTASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U61984; AAB03667.1; --
 DR HSSP; P00949; 3PMG.
 DR DictyDb; DD05092; PGM.
 DR InterPro; IPR005841; PG/PMM mutase.
 DR InterPro; IPR005844; PG_PMM_ABAT.
 DR InterPro; IPR005845; PG_PMM_ABAT.
 DR InterPro; IPR005846; PG_PMM_ABAT.
 DR InterPro; IPR005843; PG_PMM_C.
 DR Pfam; PF00408; PGM_PMM_1.
 DR Pfam; PF02878; PGM_PMM_1; 1.
 DR Pfam; PF02879; PGM_PMM_1; 1.
 DR Pfam; PF02880; PGM_PMM_1; 1.
 DR PRINTS; PR00509; PGM_PMM.
 DR PROSITE; PS00710; PGM_PMM_1.
 DR Isoemase; Phosphorylation; Magnesium.
 KW ACT_SITE 120
 FT ACT_SITE 120
 SO SEQUENCE 572 AA; 63229 MW; ED7E8F58F16F563A CRC64;

Query Match 38.1%; Score 43; DB 1; Length 572;
 Best Local Similarity 31.6%; Pred. No. 44;

Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AIDHSLIAMSNDNDYV 19
 Db 167 SYDNLGLKTYEWNDDGEFV 185

RESULT 15
 COAA_BACTU
 ID COAA_BACTU STANDARD; PRT; 674 AA.
 AC O87905;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pestidial crystal protein cry24Aa (insecticidal delta-endotoxin cryxxiva(a)) (crystalline entomocidal protoxin) (crystal protein) (insecticidal protein Jegg72) (Fragment).
 DE CRY24AA OR CRYXXIVA(A).
 GN Bacillus thuringiensis (subsp. jegashtesan).
 OS Bacillus thuringiensis (subsp. jegashtesan).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=56955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawalek M.D., Gill S.S.;
 RT "Isolation and characterization of insecticidal genes from Bacillus thuringiensis subsp. jegashtesan."
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT EPITHELIAL CELLS OF INSECTS.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U88188; AAC61891.1; --
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin_C.
 DR InterPro; IPR005639; endotoxin_N.
 DR Pfam; PF00555; endotoxin_1.
 DR Pfam; PF03944; endotoxin_C; 1.
 DR Pfam; PF03945; endotoxin_N; 1.
 DR Toxin; Sporulation.
 KW NON_TER 674
 FT NON_TER 674
 SO SEQUENCE 674 AA; 75959 MW; DA3904DAB891C978 CRC64;

Query Match 38.1%; Score 43; DB 1; Length 674;
 Best Local Similarity 46.7%; Pred. No. 53;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DRLSTLEAMSGNDND 17
 Db 138 DYLGALEAMNNKSN 152

Search completed: November 13, 2003, 09:45:32
 Job time: 8.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds

(Without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049A-13
Perfect score: 113
Sequence: 1 AIDHLSILEAMSGNDNDYVKQ 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	95.6	1338	2 Q9WXJ4	Q9WXJ4 streptococc
2	79	69.9	1575	2 Q9LCH3	Q9LCH3 streptococc
3	77	68.1	1577	2 Q54178	Q54178 streptococc
4	73	64.6	1577	2 Q55265	Q55265 streptococc
5	71	62.8	1554	2 Q8KZL5	Q8KZL5 streptococc
6	69	61.1	1477	2 Q91466	Q91466 leuconostoc
7	69	61.1	1508	2 Q9EZH5	Q9EZH5 leuconostoc
8	69	61.1	1508	2 Q52224	Q52224 leuconostoc
9	69	61.1	1512	2 Q9WXJ5	Q9WXJ5 streptococc
10	69	61.1	1518	2 Q00600	Q00600 streptococc
11	67	59.3	1449	2 Q06842	Q06842 streptococc
12	67	59.3	1449	2 Q55264	Q55264 streptococc
13	66	58.4	1016	2 Q9LC7	Q9LC7 leuconostoc
14	65	57.5	2057	2 Q9RB05	Q9RB05 leuconostoc
15	64	56.6	2835	2 Q8G9Q2	Q8G9Q2 leuconostoc
16	62	54.9	1590	2 Q59983	Q59983 streptococc

17	62	54.9	1590	2 Q55263	Q55263 streptococc
18	61	54.0	1599	2 Q00599	Q00599 streptococc
19	58	51.3	1527	2 Q9ZAR4	Q9ZAR4 leuconostoc
20	58	51.3	1527	2 Q8KRE1	Q8KRE1 leuconostoc
21	49	43.4	1428	17 Q9HSM9	Q9HSM9 halobacteri
22	49	43.4	1428	3 Q8TGS8	Q8TGS8 saccharomyc
23	49	43.4	1428	3 Q8TGS6	Q8TGS6 saccharomyc
24	49	43.4	1428	3 Q8TGS4	Q8TGS4 saccharomyc
25	49	43.4	1428	3 Q8TGS3	Q8TGS3 saccharomyc
26	49	43.4	1428	3 Q8TGS6	Q8TGS6 saccharomyc
27	49	43.4	1428	3 Q8TGS6	Q8TGS6 saccharomyc
28	49	43.4	1428	3 Q8TGS4	Q8TGS4 saccharomyc
29	49	43.4	1428	3 Q8TGS4	Q8TGS4 saccharomyc
30	48.5	42.9	576	4 Q96DX4	Q96DX4 homo sapien
31	48.5	42.9	576	6 Q95LP3	Q95LP3 macaca fasc
32	48	42.5	378	10 Q9SZH3	Q9SZH3 arabidopsi
33	48	42.5	382	16 Q8UJ06	Q8UJ06 agrobacteri
34	48	42.5	394	10 Q9AKJ2	Q9AKJ2 arabidopsi
35	48	42.5	695	16 Q9Z9G1	Q9Z9G1 chlamydia p
36	48	42.5	704	16 Q9K1Z6	Q9K1Z6 chlamydia p
37	48	42.5	956	4 Q8N1S3	Q8N1S3 homo sapien
38	48	42.5	3198	12 Q91W34	Q91W34 pea seed-bo
39	48	42.5	6885	4 Q8KXH0	Q8KXH0 homo sapien
40	47	41.6	298	5 Q19058	Q19058 caenorhabdi
41	47	41.6	769	10 Q8LAE1	Q8LAE1 arabidopsi
42	47	41.6	769	10 Q9LNU1	Q9LNU1 arabidopsi
43	47	41.6	3199	12 Q85074	Q85074 pea seed-bo
44	46	40.7	255	11 Q9CXL4	Q9CXL4 mus muscula
45	46	40.7	287	16 Q9RK11	Q9RK11 streptomyce

ALIGNMENTS

RESULT 1

Q9WXJ4 PRELIMINARY; PRT; 1338 AA.

AC Q9WXJ4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF-S.
OS Streptococcus criceti.
OC Plasmid PAM.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S.cricetus glucosyltransferase(gtfS and gtfP) genes."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB026123; BAA77236.1; -
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 10.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 95.6%; Score 108; DB 2; Length 1338;
Best Local Similarity 100.0%; Pred. No. 46-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYVK 20
Db 437 AIDHLSILEAMSGNDNDYVK 456

RESULT 2
Q9LCH3

```
ID 091CH3 PRELIMINARY; PRT; 1575 AA.
AC 091CH3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTFP.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RL encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL: AB025228; BA95201.1; -
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 17.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MM; 772A26E4D7C2E543 CRC64;

Query Match 69.9%; Score 79; DB 2; Length 1575;
Best Local Similarity 80.0%; Pred. No. 0.0019;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYK 20
DB 546 AIDHLSILEAWSGNDNDYK 565

RESULT 3
ID 054178 PRELIMINARY; PRT; 1577 AA.
AC 054178; 054247;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTFG.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA "Identification of a gene, tsg, which regulates expression of
RT glucosyltransferase and influences the Spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL: U12643; AAC43483.1; -
DR EMBL: M89776; AAA26969.1; -
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 18.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Transferase.
```

```
SQ SEQUENCE 1577 AA; 177805 MM; 5AE0328DC5E08D18 CRC64;

Query Match 68.1%; Score 77; DB 2; Length 1577;
Best Local Similarity 75.0%; Pred. No. 0.004;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYK 20
DB 548 AIDHLSILEAWSGNDNDYK 567

RESULT 4
ID 055265 PRELIMINARY; PRT; 1577 AA.
AC 055265;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTFM.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL: I35928; AAC41413.1; -
DR InterPro: IPR004829; Ceuface_antigen.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 11.
DR Pfam: PF02324; Glyco_hydro_70; 1.
DR Pfam: PD153432; Ceuface_antigen; 1.
KW Signal; Transferase.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1577 GLUCOSYLTRANSFERASE.
SQ SEQUENCE 1577 AA; 175290 MM; 3EFB898A7D3A7BE3 CRC64;

Query Match 64.6%; Score 73; DB 2; Length 1577;
Best Local Similarity 75.0%; Pred. No. 0.017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYK 20
DB 588 AIDHLSILEAWSGNDNDYK 607

RESULT 5
ID 08KZL5 PRELIMINARY; PRT; 1554 AA.
AC 08KZL5;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glucosyltransferase.
GN GTFU.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21958684; PubMed=11960691;
RA Hanada N., Fukushima K., Nomura Y., Senguku H., Hayakawa M.,
RA Mukasa H., Shiroza T., Abiko Y.;
RT "Cloning and nucleotide sequence analysis of the Streptococcus
RT sobrinus gltf gene that produces a highly branched water-soluble
RT glucan.";
```

RL Blochim. Biophys. Acta 1570:75-79 (2002).
DR EMBL; AB089438; BAC07265.1; -;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase.
SQ SEQUENCE 1554 AA; 171676 MW; 6981BCC1DAE24A73 CRC64;
Query Match 62.8%; Score 71; DB 2; Length 1554;
Best Local Similarity 72.2%; Pred. No. 0.035;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 18
DB 485 ALAHLSILEAMSLDNDQY 502
RESULT 6
ID 091466 PRELIMINARY; PRT; 1477 AA.
AC 091466;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRG.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Remaud-Simeon M., Pizzut S., Sarcabal P.,
Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternanucrase, a sucrose
glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250172; CB876565.1; -;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFCB831 CRC64;
Query Match 61.1%; Score 69; DB 2; Length 1477;
Best Local Similarity 70.0%; Pred. No. 0.069;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 20
DB 532 ANQHLSTLEDMWSHNDPEYK 551
RESULT 7
ID 09EZHS PRELIMINARY; PRT; 1508 AA.
AC 09EZHS;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase Derb742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.,
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene,"
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF294469; AAG38021.1; -;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR SEQUENCE 1508 AA; 168542 MW; E2FCFA087A84F2A CRC64;
Query Match 61.1%; Score 69; DB 2; Length 1508;
Best Local Similarity 70.0%; Pred. No. 0.07;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 20
DB 563 ANQHLSTLEDMWSHNDPEYK 582
RESULT 8
ID 052224 PRELIMINARY; PRT; 1508 AA.
AC 052224;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Remaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
synthesizing only a (1-6) glucan.";
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1; -;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CCEB57A70D1F0 CRC64;
Query Match 61.1%; Score 69; DB 2; Length 1508;
Best Local Similarity 70.0%; Pred. No. 0.07;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 20
DB 563 ANQHLSTLEDMWSHNDPEYK 582
RESULT 9
ID 09WXS PRELIMINARY; PRT; 1512 AA.
AC 09WXS;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF.
OS Streptococcus criceti.
OG Plasmid pAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S. cricetus glucosyltransferase(gtfS and gtfT) genes,"
Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB026123; BAA77237.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR00318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C6C01FC14 CRC64;

Query Match 61.1%; Score 69; DB 2; Length 1512;
 Best Local Similarity 72.2%; Pred. No. 0.071;
 Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
 |||||
 DB 485 ALAHISILEAMSGNDNDY 502

RESULT 10
 Q00600 PRELIMINARY; PRT; 1518 AA.
 AC 000600;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase I (EC 2.4.1.5) (GTF) (Dextranucrase) (Sucrose 6-
 DE Glucosyltransferase).
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=92148377; PubMed=1838391;
 RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
 RT "Molecular characterization of a cluster of at least two
 RT glucosyltransferase genes in Streptococcus salivarius ATCC 25975.";
 RL J. Gen. Microbiol. 137:2577-2593 (1991).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
 CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
 CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-
 CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- DISEASE: DENTAL CARIES.
 DR EMBL; Z11873; CA477900.1; -
 DR EMBL; M6111; AA26896.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR00318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 13.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferrase; Glycosyltransferase; Repeat; Dental caries.
 FT DOMAIN 1307 1482 6 DIRECT REPEATS.
 FT REPEAT 1307 1338 REPEAT 1.
 FT REPEAT 1338 1352 REPEAT 2.
 FT REPEAT 1372 1403 REPEAT 3.
 FT REPEAT 1404 1417 REPEAT 4.
 FT REPEAT 1437 1468 REPEAT 5.
 FT REPEAT 1469 1482 REPEAT 6.
 SQ SEQUENCE 1518 AA; 167730 MW; DAA41F717098B59A CRC64;

Query Match 61.1%; Score 69; DB 2; Length 1518;
 Best Local Similarity 66.7%; Pred. No. 0.071;
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
 |||||
 DB 501 ALAHISILEAMSGNDNDY 518

RESULT 11

068542
 ID 068542 PRELIMINARY; PRT; 1449 AA.
 AC 068542;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase N (Fragment).
 GN GTFN.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V1477;
 RA Jaffe R.I.;
 RT "Streptococcus salivarius V1477 gtfN".
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF049609; AAC05156.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR00318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferrase.
 FT NON-TER
 SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BFB CRC64;

Query Match 59.3%; Score 67; DB 2; Length 1449;
 Best Local Similarity 77.8%; Pred. No. 0.14;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
 |||||
 DB 536 AIKHLSILEAMSGNDNDY 553

RESULT 12
 Q05264 PRELIMINARY; PRT; 1449 AA.
 ID 05264;
 AC 05264;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95122197; PubMed=7822030;
 RX Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 RT coding for primer-independent glucosyltransferases.";
 RL Infect. Immun. 63:609-621 (1995).
 DR EMBL; J35495; AAC41412.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR00318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferrase.
 FT SIGNAL 1 35
 FT CHAIN 36 1449 POTENTIAL.
 SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 59.3%; Score 67; DB 2; Length 1449;
 Best Local Similarity 77.8%; Pred. No. 0.14;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
 |||||
 DB 536 AIKHLSILEAMSGNDNDY 553

RESULT 13
Q9LCJ7 PRELIMINARY; PRT; 1016 AA.
AC Q9LCJ7;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
RL mesenteroides NRRL B-512F.", 64:29-38(2000).
DR BMBJ; AB020020; BAA90527.1; -.
DR HSSP; P06278; IVOS.
DR InterPro; IPR003318; Glyco_hydro.70.
DR Pfam; PF02324; Glyco_hydro.70; 1_
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFDE13CCB47 CRC64;

Query Match 58.4%; Score 66; DB 2; Length 1016;
Best Local Similarity 70.6%; Pred. No. 0.13;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HLSTLEAWSGNDNDYK 20
DB 551 HVSILEDWSDNDAYK 567

RESULT 14
Q9RE05 PRELIMINARY; PRT; 2057 AA.
AC Q9RE05;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Alernansucrase (EC 2.4.1.140).
GN ASR.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RX MEDLINE=20080809; PubMed=10612736;
RA Arguello-Morales M.A., Remaud-Simeon M., Pizut S., Sarcabal P.,
RA Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alernansucrase, a sucrose
RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.",
RL FEMS Microbiol. Lett. 182:81-85(2000).
DR EMBL; AJ250173; CAB65910.2; -.
DR InterPro; IPR02479; CM binding.
DR InterPro; IPR03318; Glyco_hydro.70.
DR Pfam; PF01473; CM binding.1; 12.
DR Pfam; PF02324; Glyco_hydro.70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 2057 AA; 228987 MW; 62BCE938SD9A11BE CRC64;

Query Match 57.5%; Score 65; DB 2; Length 2057;
Best Local Similarity 66.7%; Pred. No. 0.43;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAWSGNDNDYK 21
DB 668 HVSILEDWSDNDAYK 685

RESULT 15
O8G9Q2 PRELIMINARY; PRT; 2835 AA.
AC O8G9Q2;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Dextranucrase (EC 2.4.1.5) (Fragment).
GN DSRE..
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22231661; PubMed=12270834;
RA Bozonnet S., Dolis-Lafargue M., Fabre E., Pizut S., Remaud-Simeon M.,
RA Mondan P., Willemot R.M.;
RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
RT synthesizing dextranucrase with two catalytic domains.",
RL J. Bacteriol. 184:5753-5761(2002).
DR EMBL; AJ430204; CAD22883.1; -..
KW Transferase; Glycosyltransferase.
FT NON TER
SQ SEQUENCE 2835 AA; 313264 MW; D03262CD735399D CRC64;

Query Match 56.6%; Score 64; DB 2; Length 2835;
Best Local Similarity 55.0%; Pred. No. 0.9;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 AIDHLSLEAWSGNDNDYK 20
DB 557 ANNHISILEDWSDNDASAYK 576

Search completed: November 13, 2003, 09:44:06
Job time : 33.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)

65.165 Million cell updates/sec

Title: US-09-290-049a-13

Perfect score: 113

Sequence: 1 AIDHSLTEAWSGNDNDYKQ 21

Scoring table: BLOSUM62

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:*

1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	66.4	545	4	US-09-604-957-4
2	75	66.4	1430	3	US-09-009-172-2
3	75	66.4	1430	3	US-09-210-361-6
4	75	66.4	1430	4	US-09-740-274-6
5	73	64.6	1475	3	US-09-007-999-2
6	73	64.6	1475	3	US-09-210-361-2
7	73	64.6	1475	4	US-09-740-274-2
8	73	64.6	1577	2	US-08-793-824-2
9	71	62.8	1375	3	US-09-210-361-4
10	71	62.8	1375	4	US-09-740-274-4
11	65	57.5	584	4	US-09-604-957-6
12	65	57.5	2057	4	US-09-499-203-2
13	58	51.3	523	4	US-09-604-957-5
14	48	42.5	661	1	US-09-198-452A-36
15	45	39.8	2366	1	US-08-480-604A-10
16	45	39.8	2366	2	US-08-405-496A-10
17	45	39.8	2366	4	US-08-915-136-10
18	45	39.8	2366	4	US-08-957-310-10
19	45	39.8	2366	4	US-10-011-366-10
20	44	38.9	270	2	US-08-852-743-5
21	44	38.9	270	3	US-09-185-370-5
22	44	38.9	487	2	US-08-712-709-8
23	44	38.9	487	3	US-09-111-444-8
24	44	38.9	487	3	US-09-541-228-8
25	44	38.9	639	2	US-08-557-309B-37
26	44	38.9	639	3	US-08-834-306-37
27	44	38.9	639	3	US-08-993-674A-37

28	44	38.9	639	4	US-09-256-976-37	Sequence 37, Appl
29	43	38.1	225	4	US-09-107-532A-4105	Sequence 4105, Ap
30	43	38.1	456	3	US-09-172-841-51	Sequence 51, Appl
31	43	38.1	456	4	US-08-951-621-51	Sequence 51, Appl
32	43	38.1	535	4	US-09-604-957-7	Sequence 7, Appl
33	43	38.1	1278	4	US-09-604-957-3	Sequence 3, Appl
34	43	38.1	487	4	US-09-685-462-8	Sequence 8, Appl
35	42	37.2	1205	4	US-09-330-330-1	Sequence 1, Appl
36	41.5	36.7	679	4	US-09-302-620B-83	Sequence 83, Appl
37	41.5	36.7	679	4	US-09-302-620B-84	Sequence 84, Appl
38	41.5	36.7	679	4	US-09-912-161-4	Sequence 4, Appl
39	41.5	36.7	679	4	US-09-912-161-6	Sequence 6, Appl
40	41	36.3	132	3	US-09-073-297-16	Sequence 16, Appl
41	41	36.3	596	4	US-09-752-165-2	Sequence 2, Appl
42	41	36.3	833	4	US-09-514-302-3	Sequence 3, Appl
43	41	36.3	1938	4	US-09-514-302-2	Sequence 2, Appl
44	40	35.4	43	1	US-08-149-839B-8	Sequence 8, Appl
45	40	35.4	43	1	US-08-451-568-8	Sequence 8, Appl

ALIGNMENTS

```
RESULT 1
US-09-604-957-4, Application US/09604957
; Sequence 4, 6486314
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJCHOUZEN, LOBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match      66.4% Score 75; DB 4; Length 545;
Best Local Similarity 75.0%; Pred. No. 0.00053;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 AIDHSLTEAWSGNDNDYK 20
      |||||
DB      75 AIDHSLTEAWSGNDNDYK 94

RESULT 2
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; PRIOR FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
```

ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 66.4%; Score 75; DB 3; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
DB 495 AINHLSILEAMSDNDPQYNK 514

RESULT 3
US-09-210-361-6

Sequence 6, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR

CURRENT APPLICATION NUMBER: US/09/210.361
EARLIER FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 66.4%; Score 75; DB 3; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
DB 495 AINHLSILEAMSDNDPQYNK 514

RESULT 4
US-09-740-274-6

Sequence 6, Application US/09740274
Patent No. 6465203
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/740.274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172

PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 66.4%; Score 75; DB 4; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
DB 495 AINHLSILEAMSDNDPQYNK 514

RESULT 5
US-09-007-999-2

Sequence 2, Application US/09007999
Patent No. 6087559
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: 0356D
CURRENT APPLICATION NUMBER: US/09/007,999
CURRENT FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 64.6%; Score 73; DB 3; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.0036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYV 19
DB 481 ANDHLSILEAMSDNDTPYL 499

RESULT 6
US-09-210-361-2

Sequence 2, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210.361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match      64.6%; Score 73; DB 3; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.0036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILKMSGNDNDYV 19
Db 481 ANDHLSILKMSGNDNDYV 499

RESULT 7
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match      64.6%; Score 73; DB 4; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.0036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILKMSGNDNDYV 19
Db 481 ANDHLSILKMSGNDNDYV 499

RESULT 8
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
```

```

; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match      64.6%; Score 73; DB 2; Length 1577;
Best Local Similarity 75.0%; Pred. No. 0.0039;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILKMSGNDNDYK 20
Db 588 AIAHLSILKMSGNDNDYK 607

RESULT 9
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRF
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match      62.8%; Score 71; DB 3; Length 1375;
Best Local Similarity 73.7%; Pred. No. 0.007;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```


Qy 1 AIDHSLTEAMSGNDNDYV 19
Db 507 ANDHSLTEAMSYNDTPYL 525

RESULT 10
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 62.8%; Score 71; DB 4; Length 1375;
Best Local Similarity 73.7%; Pred. No. 0.007;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AIDHSLTEAMSGNDNDYV 19
Db 507 ANDHSLTEAMSYNDTPYL 525

RESULT 11
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 57.5%; Score 65; DB 4; Length 584;
Best Local Similarity 66.7%; Pred. No. 0.023;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HSLTEAMSGNDNDYKQ 21
Db 78 HSLTEAMSGNDNDYKQ 95

RESULT 12
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANTZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; PRIOR FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 57.5%; Score 65; DB 4; Length 2057;
Best Local Similarity 66.7%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HSLTEAMSGNDNDYKQ 21
Db 668 HSLTEAMSGNDNDYKQ 685

RESULT 13
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 51.3%; Score 58; DB 4; Length 523;
Best Local Similarity 68.4%; Pred. No. 0.27;
Matches 13; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AIDHSLTEAMSGNDNDYV 19
Db 75 ANQHSLTEAMSGNDNDYV 93

RESULT 14
US-09-198-452A-36
; Sequence 36, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:

APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO: 36
LENGTH: 661
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-36

Query Match
Best Local Similarity 42.5%; Score 48; DB 4; Length 661;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DHLTLEAMSGNDN 16
Db 572 DHMNRWEMIGNDN 585

RESULT 15
US-08-480-604A-10
Sequence 10, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: PIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2366 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-10

Query Match
Best Local Similarity 39.8%; Score 45; DB 1; Length 2366;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IDHSLLEAMSGNDNDYV 19
Db 1733 INDLSIRYWSNDGNDPI 1750

Search completed: November 13, 2003, 09:48:01
Job time: 14.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 / Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-13

Perfect score: 113

Sequence: 1 AIDHSLTEAWSGNDNDYVKQ 21

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	75	66.4	545	10	US-09-995-749A-10	Sequence 10, App
2	75	66.4	1430	9	US-09-740-274-6	Sequence 6, App1
3	73	64.6	1475	9	US-09-740-274-2	Sequence 2, App1
4	71	62.8	1375	9	US-09-740-274-4	Sequence 4, App1
5	65	57.5	584	10	US-09-995-749A-12	Sequence 12, App
6	58	51.3	522	10	US-09-995-749A-11	Sequence 11, App
7	48.5	42.9	427	11	US-09-813-153-131	Sequence 131, App
8	45	39.8	58	9	US-09-764-887-163	Sequence 163, App
9	45	39.8	58	15	US-10-073-961-163	Sequence 163, App
10	45	39.8	2366	15	US-10-011-366-10	Sequence 10, App
11	44	38.9	207	12	US-10-225-810-28	Sequence 28, App
12	44	38.9	448	12	US-10-225-810-44	Sequence 44, App
13	44	38.9	487	9	US-09-810-808-8	Sequence 8, App1
14	44	38.9	554	11	US-09-813-408-28	Sequence 28, App
15	43	38.1	272	12	US-10-291-253A-11	Sequence 11, App

16	43	38.1	293	12	US-10-291-253A-10	Sequence 10, Appl
17	43	38.1	535	10	US-09-995-749A-13	Sequence 13, Appl
18	43	38.1	659	12	US-10-214-529-3	Sequence 3, Appl
19	43	38.1	1044	9	US-09-822-268A-5	Sequence 5, Appl
20	43	38.1	1170	9	US-09-822-268A-4	Sequence 4, Appl
21	43	38.1	1172	9	US-09-822-268A-2	Sequence 2, Appl
22	43	38.1	1237	12	US-10-154-419-82	Sequence 82, Appl
23	43	38.1	1237	12	US-10-146-733-77	Sequence 77, Appl
24	43	38.1	1237	14	US-10-024-623-32	Sequence 32, Appl
25	43	38.1	1781	10	US-09-995-749A-2	Sequence 2, Appl
26	42	37.2	116	15	US-10-078-770-196	Sequence 156, Appl
27	42	37.2	160	9	US-09-860-232A-17	Sequence 17, Appl
28	42	37.2	235	16	US-10-080-170-550	Sequence 550, Appl
29	42	37.2	240	10	US-09-738-626-6348	Sequence 5348, Appl
30	42	37.2	388	10	US-09-738-626-5242	Sequence 5242, Appl
31	42	37.2	433	14	US-10-078-829-172	Sequence 172, Appl
32	42	37.2	500	15	US-10-156-761-8558	Sequence 8558, Appl
33	42	37.2	649	15	US-10-078-770-192	Sequence 4, Appl
34	42	37.2	695	15	US-10-139-496-4	Sequence 4579, Appl
35	42	37.2	794	10	US-09-738-626-4579	Sequence 3018, Appl
36	42	37.2	867	15	US-10-128-714-3018	Sequence 8018, Appl
37	42	37.2	974	15	US-10-128-714-8018	Sequence 7, Appl
38	42	37.2	1201	12	US-10-214-529-7	Sequence 655, Appl
39	42	37.2	1915	12	US-10-238-075-655	Sequence 2, Appl
40	41.5	36.7	648	15	US-10-272-017A-2	Sequence 3, Appl
41	41.5	36.7	648	15	US-10-272-017A-3	Sequence 5, Appl
42	41.5	36.7	648	15	US-10-272-017A-5	Sequence 6, Appl
43	41.5	36.7	648	15	US-10-272-017A-6	Sequence 11, Appl
44	41.5	36.7	679	9	US-09-911-781-11	Sequence 83, Appl
45	41.5	36.7	679	11	US-09-976-800-83	

ALIGNMENTS

RESULT 1

US-09-995-749A-10

Sequence 10, Application US/0995749A

Patent No. US20020155568A1

GENERAL INFORMATION: US0995749A

APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA HENDRIKA

APPLICANT: DIRKHUIZEN, LOUBERT

APPLICANT: RAHMOUI, HAKIM

APPLICANT: LEEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

FILE REFERENCE: BO43388-CIP

CURRENT FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: US/09/995,749A

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: EPO 00201871.1

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 545

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-995-749A-10

Query Match 66.4%; Score 75; DB 10; Length 545;

Best Local Similarity 75.0%; Pred. No. 0.0034;

Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHSLTEAWSGNDNDYVK 20

DB 75 AIDHSLTEAWSGNDNDYVK 94

RESULT 2

US-09-740-274-6

Sequence 6, Application US/09740274

Patent No. US20020031826A1

```
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6
```

```
Query Match          66.4%; Score 75; DB 9; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0095;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 AIDHSLTEAMSGNDNDYK 20
         ||:|||||
Db       495 AINHSLTEAMSGNDNDPYNK 514
```

```
RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2
```

```
Query Match          64.6%; Score 73; DB 9; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.02;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 AIDHSLTEAMSGNDNDYV 19
         ||:|||||
Db       481 ANDHSLTEAMSGNDNDPYL 499
```

```
RESULT 4
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4
```

```
Query Match          62.8%; Score 71; DB 9; Length 1375;
Best Local Similarity 73.7%; Pred. No. 0.037;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 AIDHSLTEAMSGNDNDYV 19
         ||:|||||
Db       507 ANDHSLTEAMSGNDNDPYL 525
```

```
RESULT 5
US-09-995-749A-12
; Sequence 12, Application US/0995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leucorostoc mesenteroides
US-09-995-749A-12
```

```
Query Match          57.5%; Score 65; DB 10; Length 584;
Best Local Similarity 66.7%; Pred. No. 0.12;
```

Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLIEMSGNDNDYKQ 21
 |||||
 Db 78 HSLIEMSGNDNDYKQ 95

RESULT 6

US-09-995-749A-11
 ; Sequence 11, Application US/09995749A
 ; Patent No. US20020155568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
 ; APPLICANT: DUTKHUIZEN, LUBBERT
 ; APPLICANT: RAHAOUI, HAKIM
 ; APPLICANT: LEER, ROBERT-JAN
 ; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
 ; FILE REFERENCE: B043388-CIP
 ; CURRENT APPLICATION NUMBER: US/09/995,749A
 ; CURRENT FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: 09/604,957
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: EPO 00201871.1
 ; PRIOR FILING DATE: 2000-05-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 522
 ; TYPE: PRT
 ; ORGANISM: Leuconostoc mesenteroides
 US-09-995-749A-11

Query Match 51.3%; Score 58; DB 10; Length 522;
 Best Local Similarity 68.4%; Pred. No. 1.2;
 Matches 13; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AIDHSLIEMSGNDNDYV 19
 |||||
 Db 75 ANQHSLIEMSGNDNDYV 93

RESULT 7

US-09-813-153-131
 ; Sequence 131, Application US/09813153
 ; Publication No. US20030045459A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 67 Human secreted proteins
 ; FILE REFERENCE: P2023
 ; CURRENT APPLICATION NUMBER: US/09/813,153
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US/09/363,044
 ; PRIOR FILING DATE: 1999-07-29
 ; PRIOR APPLICATION NUMBER: 60/073,160
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,159
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,165
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,164
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,167
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,162
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,161
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,170
 ; PRIOR FILING DATE: 1998-01-30
 ; NUMBER OF SEQ ID NOS: 298
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 131
 ; LENGTH: 427

TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (427)
 OTHER INFORMATION: Xaa equals stop translation
 US-09-813-153-131

Query Match 42.9%; Score 48.5; DB 11; Length 427;
 Best Local Similarity 52.6%; Pred. No. 28;
 Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 DRLVLEBSWA-NDDPYLKR 21
 |||||
 Db 269 DRLVLEBSWA-NDDPYLKR 286

RESULT 8

US-09-764-887-163
 ; Sequence 163, Application US/09764887
 ; Patent No. US20020042096A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P413
 ; CURRENT APPLICATION NUMBER: US/09/764,887
 ; CURRENT FILING DATE: 2001-01-17
 ; PRIOR APPLICATION data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 658
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 163
 ; LENGTH: 58
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 NAME/KEY: SITE
 LOCATION: (48)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 LOCATION: (51)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-764-887-163

Query Match 39.8%; Score 45; DB 9; Length 58;
 Best Local Similarity 56.2%; Pred. No. 11;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AIDHSLIEMSGNDND 16
 |||||
 Db 25 AAGSLIEMSGNDND 40

RESULT 9

US-10-073-961-163
 ; Sequence 163, Application US/10073961
 ; Publication No. US20030077602A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P413C1
 ; CURRENT APPLICATION NUMBER: US/10/073,961
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: 09/764,887
 ; PRIOR FILING DATE: 2001-01-17
 ; PRIOR APPLICATION NUMBER: 60/179,065
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: 60/180,628
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: 60/214,886
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 60/217,487
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 60/225,758

;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,963
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/217,496
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,447
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/218,290
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/225,757
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/226,868
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/216,647
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,267
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/216,880
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,270
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/251,869
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/235,834
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/234,274
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/234,223
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/228,924
;; PRIOR FILING DATE: 2000-08-30
;; PRIOR APPLICATION NUMBER: 60/224,518
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/226,369
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/224,519
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,964
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/241,809
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/249,299
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/236,327
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/241,785
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/244,617
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 60/225,268
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,368
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/251,856
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/251,868
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/229,344
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/234,997
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: 60/229,343
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,345
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,513
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/231,413
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/229,509
;; PRIOR FILING DATE: 2000-09-05

;; PRIOR APPLICATION NUMBER: 60/236,367
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/236,802
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,037
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297

PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 39.8%; Score 45; DB 15; Length 58;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AIDHSLILEAWSGNDN 16
Db 25 AAGSLYLEGWSGCDH 40

RESULT 10
US-10-011-366-10

; Sequence 10, Application US/10011366
; Publication No. US20030054493A1
GENERAL INFORMATION:

APPLICANT: Williams, James A.

INVENTOR: Kink, John A.

TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPTIOPES
OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
DISEASE

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: 16-NO. US20030054493A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/957,310
FILING DATE: 23-OCT-1997
APPLICATION NUMBER: US 08/329,154
FILING DATE: 24-OCT-1994
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2366 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-011-366-10

Query Match 39.8%; Score 45; DB 15; Length 2366;
Best Local Similarity 44.4%; Pred. No. 66-02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IDHSLILEAWSGNDN 19
Db 1733 INDLSIRYVWSNDGND 1750

RESULT 11

US-10-225-810-28

; Sequence 28, Application US/10225810

; Publication No. US20030157512A1

GENERAL INFORMATION:

APPLICANT: Birmingham, Jr., John R.

TITLE OF INVENTION: Transdorins and Methods of Using Transdorin

FILE REFERENCE: McLaugh-07165

CURRENT APPLICATION NUMBER: US/10/225,810

CURRENT FILING DATE: 2002-08-21

NUMBER OF SEQ ID NOS: 76

SOFTWARE: Patent version 3.2

SEQ ID NO 28

LENGTH: 207

TYPE: PRT

ORGANISM: Homo sapiens

US-10-225-810-28

Query Match 38.9%; Score 44; DB 12; Length 207;
Best Local Similarity 61.5%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 EAWSGNDNDYVKQ 21
Db 73 ELWSDNFTDFVKQ 85

RESULT 12

US-10-225-810-44

; Sequence 44, Application US/10225810

; Publication No. US20030157512A1

GENERAL INFORMATION:

APPLICANT: Birmingham, Jr., John R.

TITLE OF INVENTION: Transdorins and Methods of Using Transdorin

FILE REFERENCE: McLaugh-07165

CURRENT APPLICATION NUMBER: US/10/225,810

CURRENT FILING DATE: 2002-08-21

NUMBER OF SEQ ID NOS: 76

SOFTWARE: PatentIn version 3.2
SEQ ID NO 44
LENGTH: 448
TYPE: PRT
ORGANISM: Mus musculus
US-10-225-810-44

Query Match 38.9%; Score 44; DB 12; Length 448;
Best Local Similarity 61.5%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 EAWSGNDNDYVKQ 21
Db 247 ELWSDNFTDFVKQ 259

RESULT 13

US-09-810-808-8
Sequence 8, Application US/09810808
Patent No. US20020042114A1
GENERAL INFORMATION:

APPLICANT: Au-Young, Janice
Guegler, Karl J.
Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/810,808
FILING DATE: 15-Mar-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/541,228
FILING DATE: <UNKNOWN>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791

US-09-810-808-8
SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 38.9%; Score 44; DB 9; Length 487;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 EAWSGNDNDYVKQ 21
Db 248 ELWSDNFTDFVKQ 260

RESULT 14

US-09-813-408-28
Sequence 28, Application US/09813408
Publication No. US20030049619A1
GENERAL INFORMATION:

APPLICANT: Delagrave, Simon
Mairs, Barry
TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial
FILE REFERENCE: HER004
CURRENT APPLICATION NUMBER: US/09/813,408
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.0
SEQ ID NO 28
LENGTH: 554
TYPE: PRT
ORGANISM: Bacillus halodurans
US-09-813-408-28

Query Match 38.9%; Score 44; DB 11; Length 554;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AIDHSLIEAWSGNDN 16
Db 450 AYDHDVVLIAASGNDN 465

RESULT 15

US-10-291-253A-11
Sequence 11, Application US/10291253A
Publication No. US20030150017A1
GENERAL INFORMATION:

APPLICANT: Botella, Jose
Graham, Michael
APPLICANT: Fairbairn, David
TITLE OF INVENTION: A Method For Facilitating Pathogen Resistance
FILE REFERENCE: nematode
CURRENT APPLICATION NUMBER: US/10/291,253A
CURRENT FILING DATE: 2003-03-31
PRIOR APPLICATION NUMBER: PR8706
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: PR8602
PRIOR FILING DATE: 2001-11-12
PRIOR APPLICATION NUMBER: US60/341404
PRIOR FILING DATE: 2001-12-14
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 272
TYPE: PRT
ORGANISM: Meloidogyne javanica
US-10-291-253A-11

Query Match 38.1%; Score 43; DB 12; Length 272;
Best Local Similarity 35.7%; Pred. No. 1.2e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 2 IDHLSIEAWSGNDNDYVKQ 21
Db 177 VDDLSIVKRWVSKSKFSNEGNDPEKQ 204

Search completed: November 13, 2003, 10:29:03
Job time : 27.4692 secs

PS Claim 13; Page 15; 29pp; Japanese.
 CC The DNA sequence from *Streptococcus sobrinus* strain 6715 encodes
 CC glucosyltransferase-I (and mutants). The DNA was obt'd. by treating
 CC *S. sobrinus* 6715 with mutanolysin, extracting the chromosomal DNA,
 CC partially digesting with Sau3AI and fractionating on agarose gel.
 CC The 3-5 kbp fragment was ligated into pUC18 and *E. coli* DM109
 CC transformed with it. A GT-I expressing clone was isolated and
 CC sequenced. The clone may be used in the development of a drug for
 CC dental caries.
 CC
 SO Sequence 1592 AA;
 Query Match 100.0%; Score 119; DB 14; Length 1592;
 Best Local Similarity 100.0%; Pred. No. 6.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAWSDNDTPYLHD 21
 Db 477 ANNHVSIVEAWSDNDTPYLHD 497
 RESULT 2
 ID AAU79285 standard; Protein; 1017 AA.
 AC AAU79285;
 DT 13-AUG-2002 (first entry)
 DE *Streptococcus mutans* monoclonal antibody-related protein #2.
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-caries; glucosyl transferase-B; immunotherapy.
 XX
 OS *Streptococcus mutans*.
 XX
 PN JP2002114709-A.
 PD 16-APR-2002.
 PF 04-OCT-2000; 2000JP-0304889.
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PR (UYN1-) UNIV NIPPON.
 XX
 PA WPI; 2002-448101/48.
 DR
 XX
 PT Anti-caries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of *Streptococcus mutans* -
 XX
 PS Claim 4; Page 17-19; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-caries agent composed of a monoclonal antibody produced by
 CC *Streptococcus mutans*, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of *Streptococcus mutans* produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a *Streptococcus mutans* monoclonal antibody-related
 CC protein.
 CC
 SO Sequence 1017 AA;
 Query Match 90.8%; Score 108; DB 23; Length 1017;
 Best Local Similarity 85.7%; Pred. No. 1.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAWSDNDTPYLHD 21

Db 447 ANNHVSIVEAWSDNDTPYLHD 467
 RESULT 3
 ID AAU98027 standard; Protein; 1475 AA.
 AC AAU98027;
 DT 27-AUG-2002 (first entry)
 DE *S. mutans* glucosyltransferase GTFB.
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture.
 XX
 OS *Streptococcus mutans*.
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0482243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 DR N-PSDB; ABR52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T,
 CC Y163N/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 XX
 SQ Sequence 1475 AA;
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSYEAMSDNDPTPLHD 21
 ||:|||||
 Db 481 ANDHLSILEAMSDNDPTPLHD 501
 RESULT 4
 AAU98030
 ID AAU98030 standard; Protein; 1475 AA.
 XX
 AC AAU98030;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant 1448V.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 448 /note="Wild-type Ile substituted by Val"
 FT
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,

CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSYEAMSDNDPTPLHD 21
 ||:|||||
 Db 481 ANDHLSILEAMSDNDPTPLHD 501
 RESULT 5
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457 /note="Wild-type Asp substituted by Asn"
 FT
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper string/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;
 SQ
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 16; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNVSVLEAVMSDNDPTLHD 21
 Db 481 ANDHLSTLEAVMSDNDPTLHD 501

RESULT 6

AAU98032
 ID AAU98032 standard; Protein; 1475 AA.
 XX
 AC AAU98032;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper string;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.

(NICH/) NICHOLS S E.

Nichols SE;

WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper string/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

[illegible]

CC	D5677_K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
CC	changes at positions from T589D, T589E, M471D, M471D/T589D, and
CC	M471D/T589E. Also included are a glucan produced by the GFP mutant,
CC	an isolated polynucleotide which encodes p1 or p2, or its complementary
CC	polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
CC	an expression cassette comprising the polynucleotide operably linked to a
CC	promoter, a vector comprising the expression cassette, host cell
CC	introduced with the vector, a transgenic plant comprising the
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC	coating composition comprising a glucan produced in a plant transformed
CC	with a gene encoding the mutant GFP, wild type or starch, a latex,
CC	thermoplastic molecule or their combinations or glucan and starch where
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line
CC	deficient in starch biosynthesis, transformed with a gene encoding a
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC	comprising the glucan (paper sizing/coating agent). The vector is useful
CC	for producing a glucan in a plant. The method comprises transforming a
CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transist sequence from
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GFP, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impact gloves to the paper during coating step.
CC	The present sequence represents a GFP mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GFPB sequence appearing as AAU98027
CC	and the information in claim 36.
CC	
XX	Sequence 1475 AA;
XX	
QY	Query Match 90.8%; Score 108; DB 23; Length 1475;
Db	Best Local Similarity 85.7%; Pred. No. 2.7e-07;
	Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
	1 ANNHVSIVEAKSDNDTPYLDH 21
	:
	481 ANDHLSITLSEAKSDNDTPYLDH 501
RESULT 8	
AAU98034	AAU98034 standard; Protein; 1475 AA.
XX	
AC	AAU98034;
XX	
DT	27-AUG-2002 (first entry).
XX	
DE	S. mutans glucosyltransferase GTFB mutant D457N/D567T.
XX	
KW	Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amyloplast; vacuole; paper manufacture; mutant; mutain.
XX	
OS	Streptococcus mutans.
OS	Synthetic.
XX	
PH	Key Location/Qualifiers
FT	Misc-difference 457
FT	/note= "Wild-type Asp substituted by Asn"
FT	Misc-difference 567
FT	/note= "Wild-type Asp substituted by Thr"
XX	
FN	US2002031826-A1.

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 EF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2,7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSGIVEAWSDNDPTFLMD 21
 ||:|||||

DB 481 ANHLSILEAWSDNDPTFLMD 501
 RESULT 9
 ID AAU98035 standard; Protein; 1475 AA.
 AC AAU98035;
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a

CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transit sequence from
CC	xylulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GFP, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GFP mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GFPB sequence appearing as AAU98027
CC	and the information in claim 36.
XX	
SQ	Sequence 1475 AA;
Query Match	90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity	85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
Df	1 ANNHVSIVEAMSDNDTPYLAMD 21 ::: ::: ::: :: 481 ANDHSLIEAMSDNDTPYLMD 501
Qy	
RESULT 10	
ID	AAU98036
XX	AAU98036 standard; Protein; 1475 AA.
XX	AAU98036;
DT	27-AUG-2002 (first entry)
XX	
DE	5. mutants glucosyltransferase GFPB mutant D567T/D571K.
XX	
KW	Glucosyltransferase; GFPB; transgenic plant; paper slating;
KV	coating composition; glucan; starch; latex; thermoplastic molecule;
KM	amyloplast; vacuole; paper manufacture; mutant; mutain.
OS	Streptococcus mutans.
XX	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT	Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FN	US2002031826-A1.
XX	
PD	14-MAR-2002.
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 98US-0007399.
PR	16-JAN-1998; 98US-0008172.
PR	20-JAN-1998; 98US-0009620.
PA	(NICH/) NICHOLS S E.
PI	Nichols SE;
WP	WPI; 2002-414332/44.

PT	Glucosyltransferase B or D protein useful for producing a glucan useful
PT	as substitutes for and additions to modified starch and latexes in
PT	paper manufacture, comprises mutations in specific positions
XX	
PS	Claim 36; Page -; 44pp; English.
XX	
CC	The invention an isolated protein comprising a glucosyltransferase
CC	(GTF) B polypeptide having changes at position from 1448V, D457N,
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC	D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC	Y159A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC	an isolated polynucleotide which encodes pl or p2, or its complementary
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC	an expression cassette comprising the polynucleotide operably linked to a
CC	promoter, a vector comprising the expression cassette, host cell
CC	introduced with the vector, a transgenic plant comprising the
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC	coating composition comprising a glucan produced in a plant transformed
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex.
CC	thermoplastic molecule or their combinations or glucan and starch where
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line
CC	deficient in starch biosynthesis, transformed with a gene encoding a
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC	comprising the glucan (paper sizing/coating agent). The vector is useful
CC	for producing a glucan in a plant. The method comprises transforming a
CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transit sequence from
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GTF, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GTFB mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GTFB sequence appearing as AA098027
CC	and the information in claim 36.
XX	
QQ	Sequence 1475 AA;
Query Match	90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity	85.7%; Pred. No. 2.7e-07.
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ANNEVSIVEAWSDNDPTYLHD 21
DB	481 ANDHLSITLAWSDNDPTYLHD 501
RESULT 11	
AAU98037	
ID	AAU98037 standard; Protein, 1475 AA.
XX	
AC	AAU98037;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	9. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
XX	
XX	Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amyloplast; vacuole; paper manufacture; mutant; mutein.
XX	
XX	Streptococcus mutans.
OS	Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substrates for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substrates for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.
 XX SQ Sequence 1475 AA;
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANHVSIVYAMSDNDPTYLHD 21
 DB 481 ANHLSIVYAMSDNDPTYLHD 501
 RESULT 12
 AAU98038
 ID AAU98038 standard; Protein, 1475 AA.
 AC AAU98038;
 XX 27-AUG-2002 (first entry)
 DT S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.
 OS Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substrates for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

RESULT 14
AAU98040
ID AAU98040 standard; Protein; 1475 AA.
XX
XX AAU98040;
XX
XX 27-AUG-2002 (first entry)
XX
XX S. mutans glucosyltransferase GTFB mutant K779Q.
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutein.
XX
XX Streptococcus mutans.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX MPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substitutes for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
XX (GTF) B polypeptide having changes at position from 1448V, D457N,
XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
XX D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
XX changes at positions from T569D, T589E, N471D, N471D/T589D, and
XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
XX an isolated polynucleotide which encodes P1 or P2, or its complementary
XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
XX an expression cassette comprising the polynucleotide operably linked to a
XX promoter, a vector comprising the expression cassette, host cell
XX introduced with the vector, a transgenic plant comprising the
XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
XX coating composition comprising a glucan produced in a plant transformed
XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
XX thermoplastic molecule or their combinations or glucan and starch where
XX the glucan is produced in the amyloplast and/or vacuole or a maize line
XX deficient in starch biosynthesis, transformed with a gene encoding a
XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper
XX comprising the glucan (paper sizing/coating agent). The vector is useful
XX for producing a glucan in a plant. The method comprises transforming a
XX plant cell with the vector, growing the plant cell under plant growing
XX conditions to produce a regenerated plant and inducing expression of the
XX polynucleotide for a time sufficient to produce the glucan in the
XX regenerated plant, where the vector contains a transit sequence from

ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
chlorophyll AB binding protein to produce a transgenic plant, and glucan
is produced in the amyloplast of potato or the vacuole of sugar beet.
Glucans are useful as substitutes for and additions to modified starch
and latexes in paper manufacture. Unlike prior art techniques, which
require input materials that produce chemical effluents, paper
manufacture utilizing the glucan produced by GTF, which utilizes
biologically produced input materials, is more cost-effective and
environmentally friendly. Moreover, glucans also exhibit thermoplastic
properties and impart gloss to the paper during coating step.
The present sequence represents a GTFB mutant of the invention.
Note: The present sequence is not shown in the specification but
was created by the indexer using the GTFB sequence appearing as AAU98027
and the information in claim 36.
SQ Sequence 1475 AA;
XX
XX Query Match 90.8%; Score 108; DB 23; Length 1475;
XX Best Local Similarity 85.7%; Freq. No. 2.7e-07;
XX Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ANNHVSIVEAMSDNDTPYLHD 21
XX |||:|||||
XX Db 481 ANDHSLIEAMSDNDTPYLHD 501
XX
XX RESULT 15
XX AAU79284
XX ID AAU79284 standard; Protein; 1476 AA.
XX
XX AAU79284;
XX
XX 13-AUG-2002 (first entry)
XX
XX Streptococcus mutans monoclonal antibody-related protein #1.
XX
XX Antibody; dental caries; water insoluble glucan synthetase;
XX anti-caries; glucosyl transferase-B; immunotherapy.
XX
XX Streptococcus mutans.
XX
XX JP2002114709-A.
XX
XX 16-APR-2002.
XX
XX 04-OCT-2000; 2000JP-0304889.
XX
XX 04-OCT-2000; 2000JP-0304889.
XX
XX (UYN1-) UNIV NIPPON.
XX
XX MPI; 2002-448101/48.
XX
XX Anti-caries agent composed of a monoclonal antibody against an
XX inhibitory enzyme against water insoluble glucan synthetase of glucosyl
XX transferase-B (GTF-B) of Streptococcus mutans -
XX
XX Claim 3; Page 13-16; 28pp; Japanese.
XX
XX The invention relates to a monoclonal antibody against dental caries and
XX an anti-caries agent composed of a monoclonal antibody produced by
XX Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
XX or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
XX inhibitory activity against water insoluble glucan synthetase of glucosyl
XX transferase-B. The monoclonal antibody specifically inhibits water
XX insoluble glucan synthetase of Streptococcus mutans produced glucosyl
XX transferase-B and is used in the immunotherapy of dental caries. This
XX sequence represents a Streptococcus mutans monoclonal antibody-related
XX protein.
XX
XX Sequence 1476 AA;
XX
XX Query Match 90.8%; Score 108; DB 23; Length 1476;

Thu Nov 13 12:12:11 2003

us-09-290-049a-14.rag

Page 11

Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 ANNHVSIVEAWSNDNDTPYLHD 21
         ||::||:|||||
Db      481 ANDHLSILEAWSNDNDTPYLHD 501
```

Search completed: November 13, 2003, 09:38:28
Job time : 39.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-14
Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 76: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1592	2 A38175	glucosyltransferase
2	108	90.8	1475	2 B31135	glbF protein precursor
3	99	83.2	1375	2 J70345	dextranase (EC
4	74	62.2	1431	2 A45866	dextranase (EC
5	68	57.1	1508	2 T31098	probable dextranase
6	67	56.3	1365	2 A41483	glucosyltransferase
7	61	51.3	1449	2 T30857	glucosyltransferase
8	61	51.3	1449	2 T30552	glucosyltransferase
9	57	47.9	1577	2 T30858	glucosyltransferase
10	56	47.1	1518	2 A44811	glucosyltransferase
11	55	46.2	1599	2 S22737	glucosyltransferase
12	53	44.5	175	2 C86205	hypothetical prote
13	49	41.2	336	1 CQBR40	BGRF2 protein - hu
14	49	41.2	524	2 D82220	conserved hypothet
15	48	40.3	331	2 B48445	glyceroldehyde-3-p
16	46	38.7	313	2 S59448	hypothetical prote
17	46	38.7	490	2 H70538	probable pdk prot
18	46	38.7	601	2 E87028	pyruvate, phosphat
19	45	38.2	418	2 D90506	4-aminobutyrate am
20	45	37.8	337	2 C64233	glyceroldehyde-3-p
21	45	37.8	347	2 T48610	hypothetical prote
22	45	37.8	525	2 T40088	RhoGEF domain cont
23	45	37.8	947	2 B86362	hypothetical prote
24	45	37.8	1384	2 T02748	hypothetical prote
25	45	37.8	4848	2 T30289	pristinamycin I sy
26	44.5	37.4	835	2 B64689	site-specific DNA-
27	44	37.0	741	2 A59414	metalloproteinase
28	44	37.0	741	2 JCS142	X-Pro dipeptidyl-p
29	44	37.0	1122	2 S64443	probable membrane

30	44	37.0	1131	2 T16217	hypothetical prote
31	44	37.0	1313	2 T29027	hypothetical prote
32	43.5	36.6	491	1 D64947	glucose-6-phosphat
33	43.5	36.6	491	2 B90949	glucose-6-phosphat
34	43.5	36.6	491	2 F85797	glucose-6-phosphat
35	43.5	36.6	491	2 AB0742	glucose-6-phosphat
36	43	36.1	78	2 S76593	hypothetical prote
37	43	36.1	194	2 H72037	conserved hypothet
38	43	36.1	194	2 C86886	CT647 hypothetical
39	43	36.1	275	2 S55978	hypothetical prote
40	43	36.1	284	2 UC7148	heat-shock sigma f
41	43	36.1	331	1 DEUTC	glyceroldehyde-3-p
42	43	36.1	346	2 E70715	hypothetical prote
43	43	36.1	351	2 D89931	hypothetical prote
44	43	36.1	396	2 T35024	probable glutathio
45	43	36.1	441	2 T36737	probable homogenit

ALIGNMENTS

RESULT 1
A38175 glucosyltransferase precursor - Streptococcus sobrinus
C:Species: Streptococcus sobrinus
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
C:Accession: A38175
R:Abn, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
U: Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A:Reference number: A38175; MUID:91123227; PMID:1704006
A:Accession: A38175
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1592 <ABO>
A:Cross-references: GB:D90213; NID:G217032; PIDN:BA14241.1; PID:d1014946; PID:G217033
A:Superfamily: cpl repeat homology
F:1093-1113/Domain: cpl repeat homology <CP1>
F:1222-1241/Domain: cpl repeat homology <CP2>
F:1287-1306/Domain: cpl repeat homology <CP3>
F:1330-1351/Domain: cpl repeat homology <CP4>
F:1352-1371/Domain: cpl repeat homology <CP5>
F:1402-1420/Domain: cpl repeat homology <CP6>
F:1465-1484/Domain: cpl repeat homology <CP7>
F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 100.0%; Score 119; DB 2; Length 1592;
Best local similarity 100.0%; Pred. No. 5.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ANNHVSIVEAMSDNDTPYLHD 21
477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 2
B31135 glbF protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C:Accession: B31135; A33128
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
U: Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the glbF gene from Streptococcus mutans.
A:Reference number: A33135; MUID:87308013; PMID:3040685
A:Accession: B31135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1475 <SHI>
A:Cross-references: GB:M7361; NID:G153639; PIDN:AA88588.1; PID:G153640
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
Submitted to the Protein Sequence Database, September 1990
A:Reference number: A33128

A/Accession: A33128
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-171,173-641,'N',643-1475 <SH2>
A/Experimental source: strain GS-5
C/Superfamily: cpl repeat homology
F/1096-1115/Domain: cpl repeat homology <CP1>
F/1224-1243/Domain: cpl repeat homology <CP2>
F/1289-1308/Domain: cpl repeat homology <CP3>
F/1354-1373/Domain: cpl repeat homology <CP4>
F/1419-1438/Domain: cpl repeat homology <CP5>

Query Match 90.8%; Score 108; DB 2; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2,6e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDPTPLHD 21
Db 481 ANDHLSILEAWSNDPTPLHD 501
|||||:|||||

RESULT 3
UT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N/Alternate names: sucrose 6-glucosyltransferase
C/Species: Streptococcus mutans
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1993
C/Accession: JT0345; C3135
R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
A/Reference number: JT0345; MUID:89137980; PMID:2576010
A/Accession: JT0345
A/Molecule type: DNA
A/Residues: 1-1375 <UED>
A/Experimental source: GS-5
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Accession: C33135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-349 <SH1>
A/Cross-references: GB:M17361
A/Genetic: gtfC
A/Function: gtfC
A/Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C/Superfamily: cpl repeat homology
C/Keywords: duplication; glycosyltransferase; hexosyltransferase
F/1-34/Domain: signal sequence #status predicted <SIG>
F/35-1375/Product: glucosyltransferase #status predicted <MAT>
F/1126-1145/Domain: cpl repeat homology <CP1>
F/1253-1272/Domain: cpl repeat homology <CP2>
F/1318-1337/Domain: cpl repeat homology <CP3>

Query Match 83.2%; Score 99; DB 2; Length 1375;
Best Local Similarity 81.0%; Pred. No. 6.1e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDPTPLHD 21
Db 507 ANDHLSILEAWSNDPTPLHD 527
|||||:|||||

RESULT 4
A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C/Accession: A45866
R/Honda, O.; Kato, C.; Kuramitsu, H.K.

J. Gen. Microbiol. 136, 2099-2105, 1990
A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
A/Reference number: A45866; MUID:91100958; PMID:2148600
A/Accession: A45866
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1431 <HON>
A/Cross-references: GB:M29296
C/Superfamily: cpl repeat homology
C/Keywords: glycosyltransferase; hexosyltransferase
F/181-201/Domain: cpl repeat homology <CP1>
F/1127-1146/Domain: cpl repeat homology <CP2>
F/1192-1211/Domain: cpl repeat homology <CP3>
F/1257-1276/Domain: cpl repeat homology <CP4>
F/1277-1297/Domain: cpl repeat homology <CP5>
F/1321-1340/Domain: cpl repeat homology <CP6>
F/1341-1361/Domain: cpl repeat homology <CP6>
F/1385-1404/Domain: cpl repeat homology <CP7>

Query Match 62.2%; Score 74; DB 2; Length 1431;
Best Local Similarity 66.7%; Pred. No. 0.005;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDPTPLHD 21
Db 495 AINHLSTLEAWSNDPTPLHD 515
|||||:|||||

RESULT 5
T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
C/Species: Leuconostoc mesenteroides
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C/Accession: T31098
R/Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A/Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A/Reference number: Z20981; MUID:98164374; PMID:9503626
A/Accession: T31098
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1508 <MON>
A/Cross-references: EMBL:AF030129; NID:92766611; PID:92766612; PIND:AA095453.1
A/Experimental source: strain NRRL B-1299
A/Genetic: dsrB
A/Function: dsrB
A/Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C/Keywords: glycosyltransferase; hexosyltransferase

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.046;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDPTPLHD 21
Db 563 ANQHLSILEAWSNDPTPLHD 583
|||||:|||||

RESULT 6
A41483
glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus
C/Species: Streptococcus sobrinus
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C/Accession: A41483
R/Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
Infect. Immun. 58, 2452-2458, 1990
A/Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltrans
A/Reference number: A41483; MUID:90316665; PMID:2142479
A/Accession: A41483
A/Molecule type: DNA
A/Residues: 1-1365 <GIL>
A/Cross-references: GB:M30943; NID:9153652; PIND:AAA6898.1; PID:9153653

C/Genetics:
A/Gene: gtfS
C/Superfamily: cpl repeat homology
C/Keywords: glycosyltransferase; hexosyltransferase

Query Match 56.3%; Score 67; DB 2; Length 1365;
Best Local Similarity 57.1%; Pred. No. 0.058;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAWSNDPTPLHD 21
Db 467 AIKHLSILEAWSNDPTPLHD 487

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius

C/Species: Streptococcus salivarius
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Accession: T30857

R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A/Reference number: Z20909; PMID:5122197; PMID:7822030

A/Accession: T30857

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1449 <SIM>

A/Cross-references: EMBL:L35495; NID:9662378; PID:9662379; PIDN:AA041412.1

C/Genetics:
A/Gene: gtfL

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAWSNDPTPLHD 21
Db 536 AIKHLSILEAWSNDPTPLHD 556

RESULT 8
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)

C/Species: Streptococcus salivarius

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C/Accession: T30552

R/Jaffe, R.I.

submitted to the EMBL Data Library, February 1998

A/Description: Streptococcus salivarius V1477 gtfN.

A/Reference number: Z20854

A/Accession: T30552

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1449 <JAF>

A/Cross-references: EMBL:AF049609; NID:92935545; PID:92935546; PIDN:AA05156.1

C/Genetics:
A/Gene: gtfN

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAWSNDPTPLHD 21
Db 536 AIKHLSILEAWSNDPTPLHD 556

RESULT 9
T30858

glucosyltransferase - Streptococcus salivarius

C/Species: Streptococcus salivarius

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C/Accession: T30858
R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A/Reference number: Z20909; PMID:5122197; PMID:7822030

A/Accession: T30858

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1577 <SIM>

A/Cross-references: EMBL:L35928; NID:9662380; PID:9662381; PIDN:AA041413.1

C/Genetics:
A/Gene: gtfm

Query Match 47.9%; Score 57; DB 2; Length 1577;
Best Local Similarity 61.1%; Pred. No. 2.5;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 HVSIVEAWSNDPTPLHD 21
Db 591 HVSIVEAWSNDPTPLHD 608

RESULT 10
A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius

C/Species: Streptococcus salivarius

C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999

C/Accession: A44811; S22726; S28809

R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A/Reference number: A44811; PMID:92148377; PMID:1838391

A/Accession: A44811

A/Molecule type: DNA

A/Residues: 1-1518 <GIF>

A/Cross-references: EMBL:Z11873; NID:947526; PIDN:CA077900.1; PID:947527

A/Note: sequence extracted from NCBI backbone (NCBI:81050, NCBI:81052)

C/Genetics:
A/Gene: gtfY

C/Superfamily: cpl repeat homology

C/Keywords: glycosyltransferase; hexosyltransferase

F/1307-1326/Domain: cpl repeat homology <CP4>

Query Match 47.1%; Score 56; DB 2; Length 1518;
Best Local Similarity 60.0%; Pred. No. 3.4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 HVSIVEAWSNDPTPLHD 18
Db 504 HVSIVEAWSNDPTPLHD 518

RESULT 11
S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C/Species: Streptococcus salivarius

C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000

C/Accession: S22737; S28810; S44811; S22727

R/Jacques, N.

submitted to the EMBL Data Library, March 1992

A/Reference number: S22726

A/Accession: S22737

A/Molecule type: DNA

A/Residues: 1-1599 <JAC>

A/Cross-references: EMBL:Z11872; NID:947530; PIDN:CA077898.1; PID:947531

A/Experimental source: ATCC 25975

R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gene

A/Reference number: A44811; PMID:92148377; PMID:1838391

A/Accession: S28810

A/Molecule type: DNA

A/Residues: 1-51 <GIF>

A:Cross-references: EMBL:Z11873

C:Genetics:

A:Gene: gtrK

C:Superfamily: cpl repeat homology

C:Keywords: glycosyltransferase; hexosyltransferase

F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 46.2%; Score 55; DB 2; Length 1599;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 HNSIVEAMSDNDTPYLHD 21
DB 494 NISILEAMSHNDPYVNE 511

RESULT 12

C66205
hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C66205

R:Theologits, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzbarg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C66205

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-175 <STO>

A:Cross-references: GB:AE005172; NID:g9954041; PIDN:AAF62215.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 44.5%; Score 53; DB 2; Length 175;
Best Local Similarity 42.1%; Pred. No. 0.78;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 NNSIVEAMSDNDTPYLHD 21
DB 109 NHOEVIDAMSDHOKPLMTD 127

RESULT 13

Q68F2 protein - human herpesvirus 4 (strain B95-8)

C:Species: human herpesvirus 4, Epstein-Barr virus

C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999

C:Accession: C43044; J01381; A03784; A03794; S33036

R:Bankier, A.T.; Deininger, P.L.; Farrell, P.D.; Barrell, B.G.

Mol. Biol. Med. 1, 21-45, 1983

A>Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus

A:Reference number: A93065; MUID:85035713; PMID:6092825

A:Accession: C43044

A:Molecule type: DNA

A:Residues: 1-336 <BAN>

A:Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24831.1; PID:g1334895

R:Bankier, A.T.; Bankier, A.T.; Blagim, M.D.; Deininger, P.L.; Farrell, P.D.; Gibson, T.J.; H

Nature 310, 207-211, 1984

A>Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A:Reference number: A03794; MUID:84270667; PMID:6087149

A:Contents: annotation; protein coding region

A:Note: neither amino acid nor nucleotide sequence is given

R:Chen, M.R.; Heu, T.Y.; Lin, S.W.; Chen, J.Y.; Yang, C.S.

J. Gen. Virol. 72, 3047-3055, 1991

A>Title: Cloning and characterization of cDNA clones corresponding to transcripts from 4

A:Reference number: J01381; MUID:92113548; PMID:1662696

A:Accession: J01381

A:Molecule type: mRNA

A:Residues: 1-336 <CHE>

A:Cross-references: GB:S77132; NID:g243314; PIDN:AA821113.1; PID:g243315

C:Superfamily: Epstein-Barr virus BGLF2 protein

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 7.1;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 ANHVSIVEAMSDNDTP 17
DB 216 AGAHVNIIRKGTEDDSP 232

RESULT 14

D82220
conserved hypothetical protein VC1268 [imported] - Vibrio cholerae (strain N16961 serogr

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: D82220

R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Base, S.; Qin, H.; Dragoi, I.; Sellers, P

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: D82220

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-524 <HEI>

A:Cross-references: GB:AE004206; GB:AE003852; NID:g9655749; PIDN:AAF94427.1; GSPDB:GN001

A:Experimental source: serogroup O1, strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1268

A:Map position: 1

Query Match 41.2%; Score 49; DB 2; Length 524;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 ANHVSIVEAMSDNDTPYLHD 21
DB 365 ANHHEIIRGWSDEVIDPALID 387

RESULT 15

B48445
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me

C:Species: Leishmania mexicana

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002

C:Accession: B48445; S25142

R:Hannett, V.; Blauw, M.; Kohl, L.; Allert, S.; Oppendoes, F.R.; Michels, P.A.M.

Mol. Biochem. Parasitol. 55, 115-126, 1992

A>Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate d

A:Reference number: A48445; MUID:93063042; PMID:1435864

A:Accession: B48445

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <HAN>

A:Cross-references: EMBL:X65220; NID:g9552; PIDN:CAA46323.1; PID:g9553

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 40.3%; Score 48; DB 2; Length 331;
Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 NNSIVEAMSDNDTPYLH 20
DB 301 NHHFVXLVSWYDNEGYSH 319

Thu Nov 13 12:12:12 2003

us-09-290-049a-14.rpr

Page 5

Search completed: November 13, 2003, 09:50:28
Job time : 14.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-14

Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLMD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1592	1	GTF2_STRDO
2	119	100.0	1597	1	GTF1_STRDO
3	108	90.8	1476	1	GTFB_STRMU
4	99	83.2	1455	1	GTFB_STRMU
5	74	62.2	1462	1	GTFD_STRMU
6	67	56.3	1365	1	GTFB_STRDO
7	49	41.2	336	1	UUI6_EBV
8	48	40.3	330	1	G3PC_LETME
9	46	38.7	313	1	YMS8_YEAST
10	45	37.8	337	1	G3P_MYCGE
11	44	37.0	569	1	SLIF_MOUSE
12	44	37.0	1122	1	YG3C_YEAST
13	44	37.0	1131	1	YPO4_CARBL
14	44	37.0	3358	1	PGCV_MOUSE
15	43.5	36.6	491	1	G6PD_ECOLI
16	43	36.1	275	1	SC35_YEAST
17	43	36.1	330	1	G3PC_TRYBB
18	43	36.1	346	1	Y943_MYCU
19	43	36.1	441	1	HGD_STRCO
20	43	36.1	603	1	POPO_HUMAN
21	43	36.1	899	1	SUHM_DROYI
22	42.5	35.7	470	1	AMP2_STRDO
23	42.5	35.7	491	1	G6PD_ERWCH
24	42	35.3	270	1	PSF1_DROME
25	42	35.3	279	1	G316_YEAST
26	42	35.3	334	1	G3P_BACST
27	42	35.3	335	1	G3P_BACCO
28	42	35.3	425	1	HGD_CAUCR
29	42	35.3	432	1	HGD_PSEAR
30	42	35.3	438	1	GATD_PYRPU
31	42	35.3	521	1	SYM_UREPA
32	42	35.3	533	1	YE09_MYCPN
33	42	35.3	775	1	AD28_HUMAN

ALIGNMENTS

34	42	35.3	776	1	AD28_MACEA	09x616 macaca fasc
35	42	35.3	1627	1	ADP1_MYCPN	P11311 mycoplasma
36	42	35.3	3698	1	P0L6_HCVB	P21530 hog cholera
37	41.5	34.9	416	1	HGD_LEGPN	Q9s4t0 legionella
38	41.5	34.9	468	1	KG3H_DICDI	P51136 dictyostell
39	41.5	34.9	578	1	VACB_YEAST	P39368 saccharomyc
40	41.5	34.9	684	1	ISHI_SCHPO	Q9y7k6 schizosach
41	41	34.5	130	1	UCR4_SCHMA	001374 schistosoma
42	41	34.5	236	1	UCR4_TOBAC	P51134 nicotiana t
43	41	34.5	241	1	HFBI_HAEIN	P35757 haemophilus
44	41	34.5	241	1	HFBI_HAEIN	P45991 haemophilus
45	41	34.5	301	1	Y166_MYCPN	P75265 mycoplasma

RESULT 1

AC	GTF2_STRDO	STANDARD;	PRT;	1592 AA.
AD	P27470;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)			
DE	(Sucrose 6-glucosyltransferase).			
OS	Streptococcus downei (Streptococcus sobrinus).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_Taxid=1317;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=6715 / Serotype G;			
RX	MEDLINE=9113227; PubMed=1704006;			
RA	Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,			
RA	Kagawa H.;			
RT	"peptide sequences for sucrose splitting and glucan binding within			
RT	Streptococcus sobrinus glucosyltransferase (water-insoluble glucan			
RT	synthetase).";			
RL	J. Bacteriol. 173:989-996(1991).			
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT			
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE			
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE			
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.			
CC	-1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-			
CC	fructose + ((1,6)-alpha-D-glucosyl)(N+1).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- DISEASE: DENTAL CARIES.			
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA			
CC	1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES			
CC	WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH			
CC	FORMS OF GLUCANS.			
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-			
CC	BINDING PROTEIN FROM S.MUTANS.			
CC	-1- SIMILARITY: Contains 16 cell wall binding repeats.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcements			
CC	or send an email to license@lsb.sib.ch).			
CC	-----			
DR	EMBL; D90213; BA14241.1;			
DR	InterPro; IPR002479; CW binding.			
DR	InterPro; IPR003318; Glyco hydro_70.			
DR	Pfam; PF01473; CW binding_1; 13.			
DR	Pfam; PF02324; Glyco hydro_70; 1.			
FT	TRANSFERASE; Glycosyltransferase; Signal; Repeat; Dental caries.			
FT	SIGNAL			
FT	CHAIN			
FT	DOMAIN			

```

FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1093 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MM; BCOA6D079351ECF CRC64;

Query Match 100.0%; Score 119; DB 1; Length 1592;
Best Local Similarity 100.0%; Pred. No. 1,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 2
GTF1_STRDO STANDARD; PRT; 1597 AA.
ID_GTF1_STRDO
AC P11001;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCB1_TaxID=1317;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFE28."
RL J. Bacteriol. 169:4271-4278 (1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17391; AAC63063.1; -.
CC InterPro: IPR002479; CM binding.
CC InterPro: IPR003318; Glyco_hydro_70.
CC Pfam: PF01473; CM binding_1_16.
CC Pfam: PF02324; Glyco_hydro_70; 1.
CC Transfaser; Glycosyltransferase; Signal; Repeat; Dental caries.
CC SIGNAL 1 38 POTENTIAL.

```

```

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT REPEAT 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1597 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MM; B9E86A200868798E CRC64;

Query Match 100.0%; Score 119; DB 1; Length 1597;
Best Local Similarity 100.0%; Pred. No. 1,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 483 ANNHVSIVEAMSDNDTPYLHD 503

RESULT 3
GTFB_STRMU STANDARD; PRT; 1476 AA.
ID_GTFB_STRMU
AC P08957; O69381; O69387; O69390; O69396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFB OR SMU1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCB1_TaxID=1309;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shirota T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans."
RL J. Bacteriol. 169:4263-4270 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans."
RT FEMS Microbiol. Lett. 161:331-336 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen."
RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.

```


	FT	VARIANT	1292	1294	DGH -> NGV (IN STRAINS GS-5, MT4467 AND MT8148).
	FT	VARIANT	1305	1369	MISSING (IN STRAIN MT4245) .
	FT	VARIANT	1326	1366	T -> V (IN STRAIN MT8148) .
	FT	VARIANT	1331	1331	I -> A (IN STRAINS GS-5 , MT4239 , MT4467 AND MT8148) .
	FT	VARIANT	1377	1377	R -> K (IN STRAIN MT8148) .
	FT	VARIANT	1398	1398	V -> I (IN STRAIN MT8148) .
	FT	VARIANT	1424	1424	D -> N (IN STRAIN MT4239) .
	FT	VARIANT	1439	1439	V -> I (IN STRAINS MT4239 AND MT8148) .
	FT	VARIANT	1444	1444	S -> P (IN STRAIN MT8148) .
	FT	CONFLICT	1455	1455	ORLVKNGVOAGKEILTEKRIGIKYDPSNCSNEAMRYVR TSSGNWYYPDNDGALISGMHVGREAVRPDEGVARYASHD QRNWMDYPDRRDGCGRSSSAVRRHSRNPDPNFPEF -> HAFLSLMTFRLEBSLSQSKVKVSNTMILLPCKPIIVYM (IN REF. 1) .
SO	SEQUENCE	1455 AA;	162965 MW;	3CB45A94EBCB6 CRC64;	
Query Match		Best Local Similarity	83.2% ; Pred.	Score 99 ; DB 1 ; Length 1455 ; No. 1.6e-07 ;	
Matches	17 ; Conservative		3 ; Mismatches	1 ; Indels	Gaps 0 ;
Oy	1 ANNHVSIVEAMSNDPTPLIAD 21	:			
Db	507 ANHDLSILEAWSYNDDPYLHD 527				
RESULT_5					
ID	GTFD_STRMU STANDARD; PRt:	1462 AA.			
AC	P4933j; O69383; O69386; O69389; O69392; O69398;				
DT	01-FEB-1996 (Rel. 33, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase) (Sucrose 6-Glucosyltransferase) .				
CN	GTDF OR SMO. 910.				
OS	Streptococcus mutans.				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OX	NCHI_TaxID=1309;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	MEDLINE=GS-5;				
RX	MEDLINE=91100958; PubMed=2148600;				
RA	Honda O., Kato C., Kuramitsu H.K.; the glucotidase sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme."				
RT	J. Gen. Microbiol. 136:2099-2105(1990).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,				
RX	MT4467 / Serotype B, and MT8148 / Serotype C;				
RX	Fujimura T., Texao Y., Hoshino T., Kawabata S., Oshima T., Sobue S., Medline=98231643; Pubmed=9570124;				
RA	Kimura S., Hamada S. ; molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans." ;				
RT	"J. Biol. Chem. 270:10555-10561 (1995)." ;				
RL	EMS Microbiol. Lett. 161:331-336(1998).				
RNA	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UA159 / ATCC 700610 / Serotype C;				
RX	MEDLINE=22295063; PubMed=12397186;				
RA	Ajdje D., Mcshan W.M., McLachlan R.E., Savic G., Chang J., Carson M.B., Primeaux C., Tian R., Kenton S., dia H., Lin S., Qian Y. Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Pereteci J.J. ; Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen " ;				
RT	Proc. Natll. Acad. Sci. U.S.A. 99:14434-14439(2002).				
RL	-I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.				

CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-S1 SYNTHESIZES BOTH FORMS OF GLUCANS.

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M29296; AAA26895.1; -

DR EMBL: D88653; BAA26103.1; -

DR EMBL: D88656; BAA26107.1; -

DR EMBL: D88659; BAA26111.1; -

DR EMBL: D88662; BAA26115.1; -

DR EMBL: D89379; BAA26121.1; -

DR EMBL: AF014932; AAN58619.1; -

DR InterPro: IPR002479; CW_binding.

DR InterPro: IPR003318; Glyco_hydro_70.

DR Pfam: PF02324; Glyco_hydro_70; 1.

KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries; Complete proteome.

KM

FT SIGNAL 1 ?

FT CHAIN ? 1462

FT DOMAIN 1332 1423

FT REPEAT 1332 1295

FT REPEAT 1296 1359

FT REPEAT 1360 1423

FT REPEAT 10 10

FT VARIANT 19 19

FT VARIANT 58 58

FT VARIANT 68 68

FT VARIANT 81 81

FT VARIANT 113 113

FT VARIANT 122 122

FT VARIANT 132 132

FT VARIANT 135 135

FT VARIANT 137 137

FT VARIANT 202 202

FT VARIANT 255 255

FT VARIANT 275 275

FT VARIANT 288 288

FT VARIANT 301 301

FT VARIANT 313 313

FT VARIANT 317 317

FT VARIANT 328 328

FT VARIANT 350 350

FT VARIANT 628 633

FT VARIANT 688 688

FT VARIANT 726 732

FT VARIANT 726 730

FT VARIANT 762 762

FT VARIANT 964 964

FT VARIANT 1019 1019

Q -> H (IN STRAIN MT4245).

D -> N (IN STRAINS MT4239 AND MT4245 AND MT4251).

E -> K (IN STRAIN MT4239).

V -> F (IN STRAIN MT4239).

F -> L (IN STRAINS MT4239, MT4251 AND MT4467).

KKKYYO -> EKEVYL (IN STRAIN MT4251).

A -> S (IN STRAIN MT4239).

IDOGSA -> ADKGNDS (IN STRAIN MT4251).

IDOGS -> ADKGN (IN STRAINS MT4239 AND MT4245).

T -> A (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

D -> Y (IN STRAIN MT4251).

E -> K (IN STRAINS MT4245 AND MT4251).

FT VARIANT 1059 1060

FT VARIANT 1060 1060

FT VARIANT 1080 1080

FT VARIANT 1142 1142

FT VARIANT 1198 1198

FT VARIANT 1220 1220

FT VARIANT 1280 1280

FT VARIANT 1282 1282

FT VARIANT 1290 1290

FT VARIANT 1311 1311

FT VARIANT 1403 1403

FT VARIANT 1425 1425

FT VARIANT 1449 1449

FT VARIANT 1449 1449

FT CONFLICT 1428 1462

FT SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;

FT SQ R (IN REF. 1).

FT RYDKNSGMVTKVTLNAGRIGIDRNGIARY -> VY

Query Match 62.2%; Score 74; DB 1; Length 1462;

Best Local Similarity 66.7%; Pred. No. 0.0015; Mismatches 5; Indels 0; Gaps 0;

Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANNHVSIWEASNDPTVYLHD 21

Db 495 AINHVSIWEASNDPTVYLHD 515

RESULT 6

GTF5_STRDO STANDARD; PRT; 1365 AA.

AC P29336;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glucosyltransferase-5 precursor (EC 2.4.1.5). (GTF-S) (Dextranucrase)

DE (sucrose 6-glucosyltransferase).

GN GTF5.

OS Streptococcus downei (Streptococcus sobrinus).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OC NCBI_TaxID=13117;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MFE28;

RX MEDLINE=90316665; PubMed=2142479;

RA Gilmore K.S., Russell R.R., Ferretti U.J.;

RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a RT glucosyltransferase that synthesizes soluble glucans."

RL Infect. Immun. 58:2452-2458 (1990).

CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDiate THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).

CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF PRIMER GLUCAN UNLIKE GTF-1.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE).

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M30943; AAA26898.1; -

DR InterPro: IPR002479; CW_binding.

DB 300 NDHFVKLVSWNDETGYSH 318

RESULT 9

YMS8_YEAST STANDARD; PRT; 313 AA.

AC Q03695;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

DR Hypothetical 35.0 kDa protein in PFK2-HPA1 intergenic region.

GN YMS206W OR YMS325.07.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SD88C / AB972;

RX PubMed=9169872;

RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,

RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,

RA Jagsels K., Iyer G., Moule S., Odell C., Pearson D., Rajandream M.A.,

RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.,

RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome

RT XIII."

RL Nature 387:90-93(1997).

CC -1 SIMILARITY: SOME, TO YEAST YMR014W.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 248755; CAA8648.1; -

DR PIR; S59448;

DR SGD; S0004819; YMR206W.

KW Hypothetical protein.

FT DOMAIN 3

FT DOMAIN 146

FT DOMAIN 149

FT DOMAIN 246

FT DOMAIN 252

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

FT DOMAIN 35017

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,

RA Nguyen D.T., Uterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,

RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Luster T.S.,

RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

RT "The minimal gene complement of Mycoplasma genitalium.";

RT Science 270:397-403(1995).

RL [2]

RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=94075230; PubMed=8253680;

RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;

RT "A survey of the Mycoplasma genitalium genome by using random

RT sequencing."

RL J. Bacteriol. 175:7918-7930(1993).

CC -1 CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -1 PATHWAY: Second phase of glycolysis; first step.

CC -1 SUBUNIT: Homotetramer (By similarity).

CC -1 SUBCELLULAR LOCATION: Cytoplasmic.

CC -1 SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U39710; AAC71523.1; -

DR EMBL; U02213; AAD12507.1; -

DR EMBL; U02178; AAD12463.1; -

DR PIR; C64233; C64233.

DR HSSP; P17721; 1HDG.

DR TIGR; MG301; -

DR InterPro; IPR000173; GAP_dhydrogenase.

DR InterPro; IPR006424; GAPDH-T.

DR Pfam; PF00044; gpdh.1.

DR Pfam; PF02800; gpdh.C.1.

DR PRINTS; PR00078; G3FDDHGNASE.

DR TIGRPFAM; TIGR01534; GAPDH-1.1.

DR PROSITE; PS00071; GAPDH; 1.

KM Glycolysis; Oxidoreductase; NAD; Complete proteome.

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157

FT BINDING 157 157


```

DR EMBL; U40935; AAA81687.1; -.
DR PIR; T16217; T16217.
DR Morneped; F31E3.4; CE01269.
DR InterPro; IPR006055; Exonuclease.
DR Pfam; PF00929; Exonuclease; 1.
DR SMART; SM00479; EXOIII; 1.
DR HYPOTHETICAL protein.
SQ SEQUENCE 1131 AA; 127174 MW; DE5A4AB0C309E7DE CRC64;

Query Match 37.0%; Score 44; DB 1; Length 1131;
Best Local Similarity 36.8%; Pred. No. 69;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NHVSIVEAMSDNPPVYLD 21
Db 315 NHGSIWVAFDRDQPOVNE 333

RESULT 14
PGCV MOUSE STANDARD; PRT; 3358 AA.
AC Q62059; Q62058; Q9CUU0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;

[1]
SEQUENCE FROM N.A. (ISOFORM V0; V1 AND V2).
RA STRAIN=C57BL/6;
RC MEDLINE=95181355; PubMed=782336;
RX Ito K., Shimomura T., Zako M., Ujita M., Kimata K.;
RA "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT proteoglycan generated by alternative splicing.";
RL J. Biol. Chem. 270:958-965(1995).

[2]
SEQUENCE FROM N.A. (ISOFORM V3).
RA STRAIN=C57BL/6;
RC MEDLINE=95181355; PubMed=782336;
RX Zako M., Shimomura T., Ujita M., Ito K., Kimata K.;
RA "Expression of pg-M(V3), an alternatively spliced form of PG-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues.";
RL J. Biol. Chem. 270:3914-3918(1995).

[3]
SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RA STRAIN=C57BL/6J; TISSUE=Skin;
RC MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Saito K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Offield T., Furuno M., Aono H., Balasubramanian R., Barth M.F.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Teyo-oka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wyshah-Bois A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).

```

```

RN [4]
RP INTERACTION WITH FBLN1.
RX PubMed=10400671;
RA Asperger A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican.";
RL J. Biol. Chem. 274:20444-20449(1999).

CC -1- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -1- SUBUNIT: Interacts with FBLN1.
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC -1- TISSUE SPECIFICITY: V2 is found only in brain.
CC -1- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -1- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----

DR EMBL; D16263; BAA03796.1; -.
DR EMBL; D28599; NOT ANNOTATED_CDS.
DR EMBL; D32040; BAA06802.1; -.
DR EMBL; AK014525; BAB29411.1; -.
DR HSSP; P01132; IEPG.
DR MGD; MGI:102889; Capg2.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR007110; IG_Like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00059; Iectin_C; 1.
DR Pfam; PF00084; Sushi; 1.
DR Pfam; PF00193; XLink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SMO0032; CCP; 1.
DR SMART; SMO0034; CLECT; 1.
DR SMART; SMO0179; EGF_CA; 1.
DR SMART; SMO0409; IG; 1.
DR SMART; SMO0445; Link; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.

```

DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS50835; IG_Like; 1.
 DR PROSITE; PS01241; LINK_2.
 KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
 KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KW Hyaluronic acid; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 3358
 FT DOMAIN 21 146
 FT DOMAIN 167 244
 FT DOMAIN 265 346
 FT DOMAIN 348 1308
 FT DOMAIN 1309 3052
 FT DOMAIN 3052 3088
 FT DOMAIN 3090 3126
 FT DOMAIN 3139 3253
 FT DOMAIN 3258 3316
 FT DOMAIN 44 130
 FT DISULFID 172 243
 FT DISULFID 196 217
 FT DISULFID 270 333
 FT DISULFID 294 315
 FT DISULFID 3056 3076
 FT DISULFID 3061 3076
 FT DISULFID 3078 3087
 FT DISULFID 3094 3105
 FT DISULFID 3099 3114
 FT DISULFID 3116 3125
 FT DISULFID 3132 3143
 FT DISULFID 3160 3252
 FT DISULFID 3228 3244
 FT DISULFID 3259 3302
 FT DISULFID 3268 3315
 FT CARBOHYD 57 57
 FT CARBOHYD 330 330
 FT CARBOHYD 351 351
 FT CARBOHYD 441 441
 FT CARBOHYD 807 807
 FT CARBOHYD 914 914
 FT CARBOHYD 951 951
 FT CARBOHYD 1305 1305
 FT CARBOHYD 1372 1372
 FT CARBOHYD 1679 1679
 FT CARBOHYD 2054 2054
 FT CARBOHYD 2244 2244
 FT CARBOHYD 2362 2362
 FT CARBOHYD 2627 2627
 FT CARBOHYD 3030 3030
 FT CARBOHYD 3332 3332
 FT CARBOHYD 3342 3342
 FT CARBOHYD 348 348
 FT VARSPPLIC 349 1308
 FT VARSPPLIC 1309 3052
 FT VARSPPLIC 349 3052
 FT CONFLICT 126 126
 FT CONFLICT 348 348
 FT CONFLICT 1658 1658
 FT CONFLICT 1674 1680
 FT CONFLICT 3358 3358
 SO SEQUENCE 3358 AA; 366938 MM; 071B80026BC0762D CRC64;

Query Match 37.0%; Score 44; DB 1; Length 3358;
 Best Local Similarity 50.0%; Pred. No. 2.6e+02;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 NH5IVEAMSDNDTPY 18
 DB 510 NH5LKELEPEKNTPT 525

RESULT 15
 ID 66PD_ECOLI STANDARD; PRT; 491 AA.
 AC P22952; P78069; O60134; O60139;
 AD 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (66PD).
 GN zwf OR B1852 OR C2265.
 OS Escherichia coli, and
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 217992;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=91123224; PubMed=1704005;
 RA Rowley D.L., Wolf R.E. Jr.;
 RT "Molecular characterization of the Escherichia coli K-12 zwf gene
 encoding glucose 6-phosphate dehydrogenase.";
 RL J. Bacteriol. 173:968-977(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Maki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
 RA Saito N., Takemoto K., Seki Y., Sivasubramanian S., Tagami H.,
 RA Takeda J., Takemoto K., Mada C., Yamamoto Y., Horinouchi T.;
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CPT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blatner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 [5]
 RP SEQUENCE OF 72-368 FROM N.A.
 RC STRAIN=various ECOL strains;
 RX MEDLINE=95064015; PubMed=7973728;
 RA Gutman D.S., Dykhuizen D.E.;
 RT "Clonal divergence in Escherichia coli as a result of recombination,
 not mutation.";
 RL Science 266:1380-1383(1994).
 [6]
 RP SEQUENCE OF 321-491 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93345818; PubMed=8344525;
 RA Carter A.T., Pearson B.M., Dickinson J.R., Lancashire W.B.;
 RT "Sequence of the Escherichia coli K-12 edd and eda genes of the
 Entner-Doudoroff pathway.";

```
RL Gene 130:155-156(1993).
CC -|- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NADP(+) = D-glucono-
CC 1,5-lactone 6-phosphate + NADPH.
CC -|- PATHWAY: Pentose phosphate pathway, first step.
CC -|- SIMILARITY: BELONGS TO THE GLUCOSE-6-PHOSPHATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M55005; AAA2475.1; -
CC EMBL; AE000279; AAC74922.1; -
CC EMBL; D90828; BAA15660.1; -
CC EMBL; AE016761; AAN80722.1; -
CC EMBL; U13783; AAA57018.1; -
CC EMBL; U13784; AAA57019.1; -
CC EMBL; U13785; AAA57020.1; -
CC EMBL; U13786; AAA57021.1; -
CC EMBL; U13787; AAA57022.1; -
CC EMBL; U13788; AAA57023.1; -
CC EMBL; U13789; AAA57024.1; -
CC EMBL; U13790; AAA57025.1; -
CC EMBL; U13791; AAA57026.1; -
CC EMBL; U13792; AAA57027.1; -
CC EMBL; U13793; AAA57028.1; -
CC EMBL; U13794; AAA57029.1; -
CC EMBL; X63694; CAA45220.1; -
CC PIR; D64947; D64947.
CC HSSP; P11411; IDPG.
CC DR SWISS-2DPAGE; P22992; COLT.
CC DR ECO2DBASE; F048_8; 6TH EDITION.
CC DR Ecogene; EG11221; zwf.
CC DR InterPro; IPR001282; G6PD.
CC DR Pfam; PF00479; G6PD_C; 1.
CC DR Pfam; PF02781; G6PD_C; 1.
CC DR PRINTS; PRO0079; G6PDHNRGNASE.
CC DR PRODOM; PD001129; G6PD; 1.
CC DR TIGRPFAMs; TIGR00871; zwf. 1.
CC DR PROSITE; PS00069; G6P_DEHYDROGENASE; 1.
CC DR Oxidoreductase; NADP; Glucose metabolism; Complete proteome.
CC KW ACT SITE 181 BY SIMILARITY.
CC FT ACT SITE 100 S -> N (IN STRAINS ECOR4 AND ECOR10).
CC FT VARIANT 100 LKSLRIDRSNVREKTVGGYTAGFA -> PEVSSPHRPLQ
CC FT CONFLICT 268 293 RTRKRRTRATYCV (IN REF. 1).
CC FT RTRKRRTRATYCV (IN REF. 1).
CC SQ SEQUENCE 491 AA; 55704 MW; 263F07D298EAFCD3 CRC64;

Query Match 36.6%; Score 43.5; DB 1; Length 491;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 6 SIVEAMS-DNDTP 17
Db 451 SITEAMAMDNDA 463
```

Search completed: November 13, 2003, 09:45:32
Job time : 7.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 31.6493 Seconds
(Without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-14
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDPTYLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1590	2	Q59983 streptococc
2	119	100.0	1590	2	Q59263 streptococc
3	76	63.9	1016	2	Q91CJ7 leucostoc
4	72	60.5	2835	2	Q869Q2 leucostoc
5	68	57.1	1477	2	Q91466 leucostoc
6	68	57.1	1508	2	Q9E2H5 leucostoc
7	68	57.1	1508	2	Q52224 leucostoc
8	68	57.1	1575	2	Q91CH3 streptococc
9	68	57.1	1577	2	Q54178 streptococc
10	67	56.3	1338	2	Q9WXJ4 streptococc
11	65	54.6	1527	2	Q92AR4 leucostoc
12	65	54.6	1527	2	Q8KRE1 leucostoc
13	61	51.3	1449	2	Q86542 streptococc
14	61	51.3	1449	2	Q55264 streptococc
15	61	51.3	1554	2	Q8KZ15 streptococc
16	60	50.4	1512	2	Q9WXJ5 streptococc

17	57	47.9	1577	2	Q55265 streptococc
18	56	47.1	1518	2	Q00600 streptococc
19	55	46.2	1599	2	Q00599 streptococc
20	55	46.2	2057	2	Q9RE05 leucostoc
21	53	44.5	175	10	Q91MJ8 arabidopsi
22	53	44.5	195	10	Q81DM8 arabidopsi
23	51	42.9	1195	10	Q06459 thermus the
24	50	42.0	601	5	Q9VGR4 drosophila
25	50	42.0	630	5	Q81A43 drosophila
26	49	41.2	524	16	Q9KSI8 vibrio chol
27	48.5	40.8	282	10	Q8RWV3 allium cepa
28	48	40.3	184	16	Q8XME4 clostridium
29	48	40.3	380	3	Q8N073 magnaporthe
30	47	39.5	193	2	Q9S445 pseudomonas
31	46.5	39.1	212	2	Q9KX16 yersinia en
32	46.5	39.1	214	11	Q8C1P7 mus musculi
33	46.5	39.1	1604	4	Q8NFA0 homo sapien
34	46	38.7	169	5	Q8S889 encephalito
35	46	38.7	271	11	Q9QZB9 rattus norv
36	46	38.7	283	5	Q9VIP7 rattus norv
37	46	38.7	283	5	Q8WSP4 drosophila
38	46	38.7	393	16	Q8Y2O6 ralstonia s
39	46	38.7	490	16	Q06579 mycobacteri
40	46	38.7	601	16	Q05566 mycobacteri
41	46	38.7	810	5	Q8T3J2 drosophila
42	46	38.7	811	5	Q9VK54 drosophila
43	46	38.7	880	5	P91643 sulfobus
44	45.5	38.2	418	17	Q97U15 streptococc
45	45	37.8	119	2	Q93RH1 streptococc

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	1590 AA.
Q59983			
ID Q59983			
AC Q59983;			
DT 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE Glucosyltransferase-I precursor (EC 2.4.1.5).			
GN GTFI.			
OS Streptococcus sobrinus.			
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC Streptococcus.			
OX NCBI_Taxid=1310;			
KX NCBI_Taxid=1310;			
RM [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=OM2176;			
RX MEDLINE=94146405; PubMed=9312602;			
RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;			
RT "DNA sequence of the glucosyltransferase gene of serotype d			
RT Streptococcus sobrinus.";			
RL DNA Seq. 4:19-27(1993).			
DR EMBL; D13858; BAA02976.1;			
DR InterPro; IPR002479; CW binding.			
DR InterPro; IPR003318; Glyco_hydro_70.			
DR Pfam; PF01473; Glyco_hydro_70; 1.			
DR Pfam; PF02324; Glyco_hydro_70; 1.			
KW Glycosyltransferase; Signal; Transferase.			
FT SIGNAL 1 38			
FT CHAIN 39 1590			
FT SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;			
QY Query Match 100.0%; Score 119; DB 2; Length 1590;			
Best Local Similarity 100.0%; Pred. No. 1.2e-09;			
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
DB 1 ANNHVSIVEAWSNDPTYLHD 21			
477 ANNHVSIVEAWSNDPTYLHD 497			

RESULT 2

Q55263 PRELIMINARY; PRT; 1590 AA.
 AC Q55263;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE GTP-I
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478."
 RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 15.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 21
 DB 477 ANNHVSIVEAWSNDPTPYLHD 497

RESULT 3

Q91CJ7 PRELIMINARY; PRT; 1016 AA.
 AC Q91CJ7;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DSRT.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 RX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512F;
 RA MEDLINE=20169623; PubMed=10705445;
 RA Funari K., Mizuno K., Takahara H., Kobayashi M.;
 RT "Gene encoding a dextranucrase-like protein in Leuconostoc mesenteroides NRRL B-512F."
 RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
 DR EMBL; AB020020; BAA90527.1; -
 DR HSP; P06278; 1VJ5
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1016 AA; 110344 MW; 8896EFD13CCCB47 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
 Best Local Similarity 61.9%; Pred. No. 0.0042;
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 21
 DB 548 ANNHVSIVEAWSNDPTPYLHD 568

RESULT 4

Q8G9Q2 PRELIMINARY; PRT; 2835 AA.
 AC Q8G9Q2;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5) (Fragment).
 GN DSRE.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 RX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=2231661; PubMed=12270834;
 RA Bozonnet S., Dolis-Lafatgue M., Fabre E., Pizut S., Renaud-Simeon M.,
 RA Morand P., Willemot R.M.;
 RT "Molecular characterisation of DSR-E, an alpha-1,2 linkage
 RT synthesising dextranucrase with two catalytic domains."
 RL J. Bacteriol. 184:5753-5761(2002).
 DR EMBL; AJ430204; CAB22883.1; -
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR GlycoSyltransferase; Transferase.
 DR SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFC831 CRC64;

Query Match 60.5%; Score 72; DB 2; Length 2835;
 Best Local Similarity 57.9%; Pred. No. 0.058;
 Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYL 19
 DB 557 ANNHVSIVEAWSNDPTPYL 575

RESULT 5

Q91A66 PRELIMINARY; PRT; 1477 AA.
 AC Q91A66;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRC.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 RX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon M., Pizut S., Sarcabal P.,
 RA Willemot R.M., Monan P.;
 RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
 RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355."
 RT Submitted (Oct-1999) to the EMBL/Genbank/DBSJ databases.
 DR EMBL; AJ250172; CAB76565.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR GlycoSyltransferase; Transferase.
 DR SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFC831 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1477;
 Best Local Similarity 57.1%; Pred. No. 0.12;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 21
 DB 532 ANNHVSIVEAWSNDPTPYLHD 552

RESULT 6
 Q9EZHS PRELIMINARY; PRT; 1508 AA.
 ID Q9EZHS

```

AC Q9BZ5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DE 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Rohyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAC38021.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; 1.E2FCFA0F87A8AF3A CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEANSDNDTPYIHD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 7
ID 052224 PRELIMINARY; PRT; 1508 AA.
AC 052224;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRL B-1299;
RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL EMBL; AF030129; AAB95453.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
RW Glycosyltransferase; transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEANSDNDTPYIHD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 8
ID Q9LCH3 PRELIMINARY; PRT; 1575 AA.
AC Q9LCH3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BA095201.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
RW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26EAD7C2E543 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1575;
Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEANSDNDTPYIHD 21
Db 546 AIKHLSILBWSNDPEYKD 566

RESULT 9
ID Q54178 PRELIMINARY; PRT; 1577 AA.
AC Q54178;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276317; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgg, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC43483.1; -.
DR EMBL; M89776; AAB26989.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
RW Transferase.
SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1577;

```

Best Local Similarity 61.8%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNDPVLHD 21
DB 548 ALKHLSTLEAWSNDNDPVLHD 568

RESULT 10

Q9WXA4 PRELIMINARY; PRT; 1338 AA.
AC Q9WXA4; 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN Streptococcus criceti.
OS Streptococcus criceti.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB026123; BAAV7236.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 56.3%; Score 67; DB 2; Length 1338;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNDPVLHD 21
DB 437 AIDHLSLEAWSNDNDPVLHD 457

RESULT 11
Q9ZAR4 PRELIMINARY; PRT; 1527 AA.
AC Q9ZAR4; 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase.
GN DEX.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512-F;
RA Bhattacharjee R., Singh D.K.S.;
RT "Cloning and Molecular Characterization of Dextranucrase Gene from
RT Leuconostoc mesenteroides NRRL B-512F";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81374; AAD10952.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW SEQUENCE 1527 AA; 169709 MW; 1DFAFA237C743398 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNDPVLHD 21
DB 581 ANQHLSLEAWSNDNDPVLHD 601

RESULT 12

Q8KRE1 PRELIMINARY; PRT; 1527 AA.
AC Q8KRE1; 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Dextranucrase Dard (EC 2.4.1.5).
GN DSRD
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RA Neubauer H., Baue A., Mollet B.;
RT "Isolation and characterization of the dextranucrase Dard of
RT Leuconostoc mesenteroides Dard";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY017384; AAG61158.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; F9P0DE220BD89668 CRC64;
SQ SEQUENCE 1527 AA; 169835 MW; F9P0DE220BD89668 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNDPVLHD 21
DB 581 ANQHLSLEAWSNDNDPVLHD 601

RESULT 13
Q68542 PRELIMINARY; PRT; 1449 AA.
AC Q68542; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 gtfN";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049609; AAC05156.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW TRANSFERASE.
FT NON TER 1449 1449
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BBB CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAMSDNDTPYLHD 21
 DB 536 AIKHLSTLEAMSHNDAYNED 556

RESULT 14
 ID Q55264 PRELIMINARY; PRT; 1449 AA.
 AC Q55264;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RI Infect. Immun. 63:609-621(1995).
 DR EMBL; L35495; AAC41412.1; -;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Signal; Transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MM; DD62F07306E86A46 CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1449;
 Best Local Similarity 57.1%; Pred. No. 1.4;
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAMSDNDTPYLHD 21
 DB 536 AIKHLSTLEAMSHNDAYNED 556

RESULT 15
 ID Q8KZL5 PRELIMINARY; PRT; 1554 AA.
 AC Q8KZL5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTF.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 sobrinus gtfV gene that produces a highly branched water-soluble
 glucan.";
 RT Blochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL; AB089438; BAC07265.1; -;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Transferase.

SQ SEQUENCE 1554 AA; 171676 MM; 6981BCC1DAE24A73 CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1554;
 Best Local Similarity 61.1%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 HVSTVEAMSDNDTPYLHD 21
 DB 488 HSTLEAMSLNDQYNE 505

Search completed: November 13, 2003, 09:44:06
 Job time : 31.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds

(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-14

Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLMD 21Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	90.8	1475	3	US-09-007-999-2
2	108	90.8	1475	3	US-09-210-361-2
3	108	90.8	1475	4	US-09-740-274-2
4	99	83.2	1375	3	US-09-210-361-4
5	99	83.2	1375	4	US-09-740-274-4
6	74	62.2	1430	3	US-09-604-957-4
7	74	62.2	1430	3	US-09-008-172-2
8	74	62.2	1430	3	US-09-210-361-6
9	74	62.2	1430	4	US-09-740-274-6
10	65	54.6	1577	2	US-09-604-957-5
11	57	47.9	1577	2	US-08-793-824-2
12	55	46.2	584	4	US-09-604-957-6
13	55	46.2	2057	4	US-09-499-203-2
14	48	40.3	535	4	US-09-604-957-7
15	48	40.3	1278	4	US-09-604-957-3
16	47	39.5	792	4	US-09-562-737-96
17	43.5	36.6	1092	4	US-09-275-608-3
18	43	36.1	151	4	US-09-198-452A-815
19	43	36.1	234	4	US-09-134-001C-4514
20	43	36.1	647	4	US-09-134-001C-5458
21	42	35.3	200	2	US-08-836-442-4
22	42	35.3	287	4	US-09-252-991A-31548
23	42	35.3	327	4	US-09-107-532A-6181
24	42	35.3	334	6	5290690-11
25	42	35.3	403	4	US-09-252-991A-27960
26	42	35.3	454	4	US-09-252-991A-28000
27	42	35.3	455	3	US-09-362-473-6

28	42	35.3	529	2	US-08-836-442-3	Sequence 3, Appli
29	42	35.3	1627	1	US-07-665-792E-9	Sequence 9, Appli
30	42	35.3	3898	4	US-08-750-717-2	Sequence 2, Appli
31	41.5	34.9	69	4	US-09-252-991A-17383	Sequence 17383, A
32	41.5	34.9	180	4	US-09-328-352-4952	Sequence 4952, Ap
33	41	34.5	88	4	US-09-252-991A-21480	Sequence 21480, A
34	41	34.5	259	1	US-08-277-231A-3	Sequence 3, Appli
35	41	34.5	259	2	US-08-477-326-6	Sequence 6, Appli
36	41	34.5	259	2	US-08-596-291-3	Sequence 3, Appli
37	41	34.5	2465	3	US-09-100-804-3	Sequence 3, Appli
38	41	34.5	2465	3	US-09-080-855-12	Sequence 12, Appli
39	41	34.5	2466	3	US-09-566-076-12	Sequence 12, Appli
40	41	34.5	2466	5	PCT-US94-09943-2	Sequence 2, Appli
41	41	34.5	2466	5	US-09-290-640-46	Sequence 46, Appli
42	41	34.5	2485	2	US-08-876-991-2	Sequence 2, Appli
43	41	34.5	3898	2	US-09-853-72	Sequence 2, Appli
44	41	34.5	3898	2	US-09-214-278-1	Sequence 1, Appli
45	40.5	34.0	214	3	US-09-214-278-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-007-999-2

; Sequence 2, Application US/09007999

; Patent No. 6087559

; GENERAL INFORMATION:

Query Match 90.8%; Score 108; DB 3; Length 1475;

Best Local Similarity 85.7%; Pred. No. 1.3e-06;

Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ANNHVSIVEAMSDNDTPYLMD 21
|||:|||||
Db 481 ANNHVSIVEAMSDNDTPYLMD 501

RESULT 2
US-09-210-361-2

; Sequence 2, Application US/09210361

; Patent No. 6284479

; GENERAL INFORMATION:

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-210-361-2

Query Match          90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
DB      481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-2

Query Match          90.8%; Score 108; DB 4; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
DB      481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; EARLIER APPLICATION NUMBER: 08/482,711
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-4

Query Match          83.2%; Score 99; DB 3; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
DB      507 ANDHLSILEAMSYNDTPYLHD 527

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-4

Query Match          83.2%; Score 99; DB 4; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
DB      507 ANDHLSILEAMSYNDTPYLHD 527

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6466314
```

```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-604-957-4

Query Match          62.2%; Score 74; DB 4; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.0098;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIVKASNDNTPYLHD 21
Db 75 AINHLSILEAWSNDNPQYKND 95

RESULT 7
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-008-172-2

Query Match          62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIVKASNDNTPYLHD 21
Db 495 AINHLSILEAWSNDNPQYKND 515

RESULT 8
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
```

```

; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-210-361-6

Query Match          62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIVKASNDNTPYLHD 21
Db 495 AINHLSILEAWSNDNPQYKND 515

RESULT 9
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-6

Query Match          62.2%; Score 74; DB 4; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIVKASNDNTPYLHD 21
Db 495 AINHLSILEAWSNDNPQYKND 515

RESULT 10
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
```

APPLICANT: RAHAOUI, HAKIM
 APPLICANT: LEER, ROBERT-JAN
 TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
 FILE REFERENCE: BO 43388
 CURRENT APPLICATION NUMBER: US/09/604,957
 CURRENT FILING DATE: 2000-06-28
 PRIOR APPLICATION NUMBER: 00201871.1
 PRIOR FILING DATE: 2000-05-25
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 5
 LENGTH: 523
 TYPE: PRT
 ORGANISM: Leuconostoc mesenteroides
 US-09-604-957-5

Query Match 54.6%; Score 65; DB 4; Length 523;
 Best Local Similarity 57.1%; Pred. No. 0.025;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 DB 75 ANQHLSTLEDMWNGKDPQYVN 95

RESULT 11
 US-08-793-824-2
 Sequence 2, Application US/08793824

PATENT INFORMATION:
 PATENT NO. 5981838
 APPLICANT: Simpson, Christine Lynn
 APPLICANT: Giffard, Philip Morrison
 APPLICANT: Jacques, Nicholas Anthony
 TITLE OF INVENTION: Genetic Manipulation of plants to
 INCREASE STORED CARBOHYDRATES
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESSES:
 ADDRESSER: Griffith Hack & Co
 STREET: Level 8, 168 Walker Street
 CITY: No. 5981838th Sydney
 STATE: New South Wales
 COUNTRY: Australia
 ZIP: 2060

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/793,824
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: AU PM7643
 FILING DATE: 24-AUG-1994
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 61 2 9957 5944
 TELEFAX: 61 2 957 6288
 TELEX: 26547

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1577 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Streptococcus salivarius
 US-08-793-824-2

Query Match 47.9%; Score 57; DB 2; Length 1577;
 Best Local Similarity 61.1%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 HVSIVEAMSDNDTPYLHD 21
 DB 591 HSLTLEAMSYNDHQYKND 608

RESULT 12
 US-09-604-957-6
 Sequence 6, Application US/09604957
 Patent No. 6486314

GENERAL INFORMATION:
 APPLICANT: VAN GELT-SCHUTTEN, GERITDINA HENDRIKA
 APPLICANT: DIJKHOIZEN, LUBBERT
 APPLICANT: RAHAOUI, HAKIM
 APPLICANT: LEER, ROBERT-JAN
 TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
 FILE REFERENCE: BO 43388
 CURRENT APPLICATION NUMBER: US/09/604,957
 CURRENT FILING DATE: 2000-06-28
 PRIOR APPLICATION NUMBER: 00201871.1
 PRIOR FILING DATE: 2000-05-25
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 6
 LENGTH: 584
 TYPE: PRT
 ORGANISM: Leuconostoc mesenteroides
 US-09-604-957-6

Query Match 46.2%; Score 55; DB 4; Length 584;
 Best Local Similarity 45.0%; Pred. No. 1.1;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
 DB 75 ANKHLSTLEDMWNGKDPQYVN 94

RESULT 13
 US-09-499-203-2
 Sequence 2, Application US/09499203
 Patent No. 6570065

GENERAL INFORMATION:
 APPLICANT: KOSKMAN, Jens
 APPLICANT: WELSH, Thomas
 APPLICANT: QUANZ, Martin
 APPLICANT: KNUTH, Karola
 TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
 FILE REFERENCE: 147-196P
 CURRENT APPLICATION NUMBER: US/09/499,203
 CURRENT FILING DATE: 2000-02-08
 NUMBER OF SEQ ID NOS: 54
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 2
 LENGTH: 2057
 TYPE: PRT
 ORGANISM: Leuconostoc mesenteroides
 US-09-499-203-2

Query Match 46.2%; Score 55; DB 4; Length 2057;
 Best Local Similarity 45.0%; Pred. No. 4.5;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
 DB 665 ANKHLSTLEDMWNGKDPQYVN 684

RESULT 14
 US-09-604-957-7
 Sequence 7, Application US/09604957
 Patent No. 6486314
 GENERAL INFORMATION:

```

1  APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
2  APPLICANT: DIJKHUIZEN, LUBBERT
3  APPLICANT: RAHMOUI, HAKIM
4  APPLICANT: LEER, ROBERT-JAN
5  TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
6  FILE REFERENCE: BO 43388
7  CURRENT APPLICATION NUMBER: US/09/604,957
8  CURRENT FILING DATE: 2000-06-28
9  PRIOR APPLICATION NUMBER: 00201871.1
10 PRIOR FILING DATE: 2000-05-25
11 NUMBER OF SEQ ID NOS: 17
12 SOFTWARE: PatentIn Ver. 2.1
13 SEQ ID NO 7
14 LENGTH: 535
15 TYPE: PRN
16 ORGANISM: Lactobacillus reuteri
17 US-09-604-957-7

```

Query Match	40.3%	Score 48	DB 4	Length 535
Best Local Similarity	35.0%	Pred. No. 12		
Matches	7	Conservative	6	Mismatches 7
				Indels 0
				Gaps 0

QY	1	ANNHVSIVEAWSNDNTPYLH	20
		: :: : : : :	
Db	74	SNKHINILEDWNHADPEYFN	93

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHOUTEN, GERRITJINA HENDRIKA
; APPLICANT: DIJKHUIZEN, IJUBERT
; APPLICANT: RAHAOUT, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRF
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

```

Query Match	40.3%	Score 48;	DB 4;	Length 1278;
Best Local Similarity	35.0%	Pred. No. 34;		
Matches	7;	Conservative	6;	Mismatches 7;
				Indels 0;
				Gaps 0;

```
QY      1 ANNHVSIVEAWSNDNTPYLH 20  
       :|::||:|:|:  
Db     550 SNKHINILEDWNHADPEYFN 569
```

Search completed: November 13, 2003, 09:48:01
Job time : 13.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 / Search time 27.4692 seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-14
Sequence: 119
1 ANNHVSIVEAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	90.8	1475	9 US-09-740-274-2	Sequence 2, Appli
2	99	83.2	1375	9 US-09-740-274-4	Sequence 4, Appli
3	74	62.2	545	10 US-09-995-749A-10	Sequence 10, Appli
4	74	62.2	1430	9 US-09-740-274-6	Sequence 6, Appli
5	65	54.6	522	10 US-09-995-749A-11	Sequence 11, Appli
6	55	46.2	584	10 US-09-995-749A-12	Sequence 12, Appli
7	48	40.3	535	10 US-09-995-749A-13	Sequence 13, Appli
8	48	40.3	1781	10 US-09-995-749A-2	Sequence 2, Appli
9	47	39.5	749	15 US-10-211-962-96	Sequence 96, Appli
10	46.5	39.1	1604	9 US-09-888-615-73	Sequence 73, Appli
11	46	38.7	502	12 US-10-009-823A-1	Sequence 1, Appli
12	46	38.7	1494	15 US-10-195-144-17	Sequence 17, Appli
13	45	37.8	199	10 US-09-738-626-4571	Sequence 4571, Ap
14	45	37.8	665	15 US-10-156-761-14923	Sequence 14923, A
15	45	37.8	972	10 US-09-924-154-16	Sequence 16, Appli

16	44	37.0	15	12	US-10-295-693-90	Sequence 90, Appli
17	44	37.0	15	15	US-10-067-649-90	Sequence 90, Appli
18	43.5	36.6	202	10	US-09-280-197-11	Sequence 11, Appli
19	43.5	36.6	202	12	US-10-448-139-11	Sequence 11, Appli
20	43.5	36.6	491	9	US-09-815-242-10190	Sequence 10190, A
21	43.5	36.6	491	9	US-09-815-242-13756	Sequence 13756, A
22	43.5	36.6	1092	10	US-09-423-126-5	Sequence 5, Appli
23	43	36.1	27	15	US-10-093-908-26	Sequence 26, Appli
24	43	36.1	67	15	US-10-050-704-324	Sequence 324, App
25	43	36.1	87	15	US-10-050-704-149	Sequence 149, App
26	43	36.1	332	10	US-09-454-279-20	Sequence 20, Appli
27	43	36.1	437	11	US-09-374-046A-66	Sequence 66, Appli
28	43	36.1	539	11	US-10-230-026-34	Sequence 34, Appli
29	43	36.1	603	15	US-10-050-678-2	Sequence 2, Appli
30	43	36.1	603	15	US-10-092-908-35	Sequence 35, Appli
31	43	36.1	715	15	US-10-156-761-7562	Sequence 7562, Ap
32	42	35.3	349	12	US-10-220-381-22	Sequence 22, Appli
33	42	35.3	432	9	US-09-815-242-5096	Sequence 5096, Ap
34	42	35.3	455	9	US-09-742-954-6	Sequence 6, Appli
35	42	35.3	7349	12	US-10-314-657-46	Sequence 46, Appli
36	41.5	34.9	49	12	US-10-029-386-28981	Sequence 28981, A
37	41	34.5	120	14	US-10-001-835-146	Sequence 146, App
38	41	34.5	302	9	US-09-815-242-14004	Sequence 9, Appli
39	41	34.5	358	11	US-09-095-478-7	Sequence 7, Appli
40	41	34.5	381	11	US-09-095-478-7	Sequence 7, Appli
41	41	34.5	438	15	US-10-156-761-14114	Sequence 14114, A
42	41	34.5	749	15	US-10-156-761-11729	Sequence 11729, A
43	41	34.5	1267	15	US-10-059-585-56	Sequence 56, Appli
44	41	34.5	2466	12	US-10-177-980-12	Sequence 12, Appli
45	41	34.5	2485	9	US-09-802-669-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OR INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2
Query Match 90.8%; Score 108; DB 9; Length 1475;
Best Local Similarity 85.7%; Pred. No 7.3e-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
1 ANNHVSIVEAMSDNDTPYLHD 21
|||:|||||

Db 481 ANDHLSIEAMSDNDPTPLMD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/110,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match

Best Local Similarity 83.2%; Score 99; DB 9; Length 1375;
Best Local Similarity 81.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDPTPLMD 21

Db 507 ANDHLSIEAMSDNDPTPLMD 527

RESULT 3

US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match

Best Local Similarity 62.2%; Score 74; DB 10; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDPTPLMD 21

Db 75 AINHLSIEAMSDNDPQYNKD 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match

Best Local Similarity 62.2%; Score 74; DB 9; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.012; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDPTPLMD 21

Db 495 AINHLSIEAMSDNDPQYNKD 515

RESULT 5

US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match

Best Local Similarity 54.6%; Score 65; DB 10; Length 522;
Best Local Similarity 57.1%; Pred. No. 0.097;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 21
 Db 75 ANKHSTILEDWMDPLVYTD 95

RESULT 6
 US-09-995-749A-12
 ; Sequence 12, Application US/09995749A
 ; Patent No. US20020155568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
 ; APPLICANT: DIJKHUIZEN, LUBBERT
 ; APPLICANT: RAHAOUI, HAKIM
 ; APPLICANT: LEER, ROBERT-JAN
 ; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
 ; FILE REFERENCE: BO43388-CIP
 ; CURRENT APPLICATION NUMBER: US/09/995,749A
 ; CURRENT FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: 09/604,957
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: EPO 00201871.1
 ; PRIOR FILING DATE: 2000-05-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 584
 ; TYPE: PRT
 ; ORGANISM: Leuconostoc mesenteroides
 US-09-995-749A-12

Query Match 46.2%; Score 55; DB 10; Length 584;
 Best Local Similarity 45.0%; Pred. No. 3.8;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
 Db 75 ANKHSTILEDWMDPLVYTD 94

RESULT 7
 US-09-995-749A-13
 ; Sequence 13, Application US/09995749A
 ; Patent No. US20020155568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
 ; APPLICANT: DIJKHUIZEN, LUBBERT
 ; APPLICANT: RAHAOUI, HAKIM
 ; APPLICANT: LEER, ROBERT-JAN
 ; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
 ; FILE REFERENCE: BO43388-CIP
 ; CURRENT APPLICATION NUMBER: US/09/995,749A
 ; CURRENT FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: 09/604,957
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: EPO 00201871.1
 ; PRIOR FILING DATE: 2000-05-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 535
 ; TYPE: PRT
 ; ORGANISM: Lactobacillus reuteri
 US-09-995-749A-13

Query Match 40.3%; Score 48; DB 10; Length 535;
 Best Local Similarity 35.0%; Pred. No. 4.1;
 Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
 Db 74 SKHINILEDWMDPLVYTD 93

RESULT 8
 US-09-995-749A-2
 ; Sequence 2, Application US/09995749A
 ; Patent No. US20020155568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
 ; APPLICANT: DIJKHUIZEN, LUBBERT
 ; APPLICANT: RAHAOUI, HAKIM
 ; APPLICANT: LEER, ROBERT-JAN
 ; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
 ; FILE REFERENCE: BO43388-CIP
 ; CURRENT APPLICATION NUMBER: US/09/995,749A
 ; CURRENT FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: 09/604,957
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: EPO 00201871.1
 ; PRIOR FILING DATE: 2000-05-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1781
 ; TYPE: PRT
 ; ORGANISM: Lactobacillus reuteri
 US-09-995-749A-2

Query Match 40.3%; Score 48; DB 10; Length 1781;
 Best Local Similarity 35.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
 Db 1053 SKHINILEDWMDPLVYTD 1072

RESULT 9
 US-10-211-962-96
 ; Sequence 96, Application US/10211962
 ; Publication No. US20030082640A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hertz, Joachim
 ; APPLICANT: Gotthardt, Michael
 ; TITLE OF INVENTION: LDL Receptor Signaling Pathways
 ; FILE REFERENCE: UTSW0708
 ; CURRENT APPLICATION NUMBER: US/10/211,962
 ; CURRENT FILING DATE: 2002-08-01
 ; PRIOR APPLICATION NUMBER: US/09/562,737
 ; PRIOR FILING DATE: 2000-05-01
 ; NUMBER OF SEQ ID NOS: 132
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 96
 ; LENGTH: 749
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Sequence
 US-10-211-962-96

Query Match 39.5%; Score 47; DB 15; Length 749;
 Best Local Similarity 53.3%; Pred. No. 86;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 STVAMSDNDTPYLH 20
 Db 136 SAVBWTDSATPHH 150

RESULT 10
 US-09-888-615-73
 ; Sequence 73, Application US/09888615
 ; Patent No. US20020064856A1

GENERAL INFORMATION:
 APPLICANT: PLOMMAN, GREGORY
 APPLICANT: WHITE, DAVID
 APPLICANT: CAENEPEEL, SEAN
 APPLICANT: CHARVOCZAK, GLEN
 APPLICANT: MANNING, GERARD
 APPLICANT: SUDARSANAM, SUCHA
 TITLE OF INVENTION: NOVEL PROTEASES
 FILE REFERENCE: 038602/1214
 CURRENT APPLICATION NUMBER: US/09/888,615
 CURRENT FILING DATE: 2001-06-26
 PRIOR APPLICATION NUMBER: 60/214,047
 PRIOR FILING DATE: 2000-06-26
 NUMBER OF SEQ ID NOS: 150
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 73
 LENGTH: 1604
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-888-615-73

Query Match 39.1%; Score 46.5; DB 9; Length 1604;
 Best Local Similarity 47.4%; Pred. No. 2.4e+02;
 Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 5 VSIVEAMSDN--DTPYH 20
 DB 294 VALLEWKMDRTDPELH 312

RESULT 11
 US-10-009-823A-1
 Sequence 1, Application US/10009823A
 Publication No. US20030157120A1

GENERAL INFORMATION:
 APPLICANT: Panaccio, Michael
 APPLICANT: Rosey, Everett Lee
 APPLICANT: Sinistaj, Meri
 APPLICANT: Hasse, Declaf
 APPLICANT: Parsons, Jim
 APPLICANT: Ankenbauer, Robert G.
 TITLE OF INVENTION: LAWSONIA DERIVED GENE AND RELATED FLGE
 TITLE OF INVENTION: POLYPEPTIDES, PEPTIDES AND PROTEINS AND THEIR USES
 FILE REFERENCE: DAV150.001APC
 CURRENT APPLICATION NUMBER: US/10/009,823A
 CURRENT FILING DATE: 2002-08-13
 PRIOR APPLICATION NUMBER: PCT/AU00/00437
 PRIOR FILING DATE: 2000-05-11
 PRIOR APPLICATION NUMBER: US 60/133,973
 PRIOR FILING DATE: 1999-05-13
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 502
 TYPE: PRT
 ORGANISM: Lawsonia intracellularis
 US-10-009-823A-1

Query Match 38.7%; Score 46; DB 12; Length 502;
 Best Local Similarity 41.2%; Pred. No. 79;
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNTP 17
 DB 185 ANPYFALBSWKNGTIP 201

RESULT 12
 US-10-195-144-17
 Sequence 17, Application US/10195144
 Publication No. US20030126646A1
 GENERAL INFORMATION:
 APPLICANT: BROWN, GREGORY G.

APPLICANT: FORMANOVA, NATASA
 APPLICANT: DENDY, CHARLES
 APPLICANT: LANDRY, BENOT S.
 APPLICANT: CHEUNG, WING
 APPLICANT: JIN, HUA
 TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
 FILE REFERENCE: 16313-0136
 CURRENT APPLICATION NUMBER: US/10/195,144
 CURRENT FILING DATE: 2002-10-01
 PRIOR APPLICATION NUMBER: 60/305,026
 PRIOR FILING DATE: 2001-07-12
 PRIOR APPLICATION NUMBER: 60/305,363
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 60/308,736
 PRIOR FILING DATE: 2001-07-30
 NUMBER OF SEQ ID NOS: 128
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 17
 LENGTH: 1494
 TYPE: PRT
 ORGANISM: Raphanus sativum
 US-10-195-144-17

Query Match 39.7%; Score 46; DB 15; Length 1494;
 Best Local Similarity 41.2%; Pred. No. 2.7e+02;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHVSIVEAMSDNTPYL 19
 DB 1130 NTISLAEITWLNSEPF 1146

RESULT 13
 US-09-738-626-4571
 Sequence 4571, Application US/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 4571
 LENGTH: 199
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4571

Query Match 37.8%; Score 45; DB 10; Length 199;
 Best Local Similarity 58.3%; Pred. No. 40;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 IVEAMSDNTPY 18
 DB 177 LIVEWTDQDTRY 188

RESULT 14
 US-10-156-761-14923
 ; Sequence 14923, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMIURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 14923
 ; LENGTH: 665
 ; TYPE: PRY
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-14923

Query Match 37.8%; Score 45; DB 15; Length 665;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHVSIVEAMSDNDPTYLH 20
 DB 88 NWTYLVVAMSDPTVTRH 105

RESULT 15
 US-09-924-154-16
 ; Sequence 16, Application US/09924154
 ; Patent No. US20020127241A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nam, David L.
 ; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
 ; FILE REFERENCE: 05213-0465 43170-262105
 ; CURRENT APPLICATION NUMBER: US/09/924,154
 ; CURRENT FILING DATE: 2001-08-07
 ; PRIOR APPLICATION NUMBER: US 60/223,525
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 972
 ; TYPE: PRY
 ; ORGANISM: Mammalian
 US-09-924-154-16

Query Match 37.8%; Score 45; DB 10; Length 972;
 Best Local Similarity 44.4%; Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 NHVSIVEAMSDNDPTYLH 20
 DB 304 NHTGVCLMEDDDNDQYLH 321

Search completed: November 13, 2003, 10:29:04
 Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 40.3507 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049a-15
Perfect score: 110
Sequence: 1 VPSYSPIRAHSEVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	1375	23 AAU98028	S. mutans glucosyl
2	110	100.0	1375	23 AAU98028	S. mutans glucosyl
3	105	95.5	1017	23 AAU98028	Streptococcus muta
4	105	95.5	1476	23 AAU98028	Streptococcus muta
5	99	90.0	1475	23 AAU98027	S. mutans glucosyl
6	99	90.0	1475	23 AAU98030	S. mutans glucosyl
7	99	90.0	1475	23 AAU98031	S. mutans glucosyl
8	99	90.0	1475	23 AAU98033	S. mutans glucosyl
9	99	90.0	1475	23 AAU98033	S. mutans glucosyl

10	99	90.0	1475	23 AAU98040	S. mutans glucosyl
11	98	89.1	1475	23 AAU98035	S. mutans glucosyl
12	98	89.1	1592	14 AAU98032	Glucosyltransferase
13	92	83.6	1475	23 AAU98032	S. mutans glucosyl
14	92	83.6	1475	23 AAU98034	S. mutans glucosyl
15	91	82.7	1475	23 AAU98036	S. mutans glucosyl
16	91	82.7	1475	23 AAU98037	S. mutans glucosyl
17	91	82.7	1475	23 AAU98038	S. mutans GTP mut
18	82	74.5	1527	23 AAU80055	Leuconostoc mesent
19	78	70.9	2835	23 AAU80055	Dextran saccharase
20	76	69.1	15	23 ABB98574	Dextrane-saccharas
21	75	68.2	1781	23 AAU74519	Lactobacillus reut
22	75	68.2	1430	23 AAU98044	S. mutans glucosyl
23	75	68.2	1430	23 AAU98044	Dextrane-saccharas
24	71	64.5	15	23 ABB98651	S. mutans glucosyl
25	71	64.5	1430	23 AAU98042	S. mutans glucosyl
26	71	64.5	1430	23 AAU98045	L. mesenteroides a
27	71	64.5	2057	21 AAB10667	Alpha-D-glucosyltr
28	70	63.6	1577	17 AAR91047	S. mutans glucosyl
29	68	61.8	1430	23 AAU98029	S. mutans glucosyl
30	68	61.8	1430	23 AAU98043	Dextrane-saccharas
31	67	60.9	15	23 ABB98654	Dextrane-saccharas
32	65	59.1	15	23 ABB98653	Dextrane-saccharas
33	65	59.1	15	23 ABB98655	Dextrane-saccharas
34	63	57.3	15	23 ABB98579	Dextrane-saccharas
35	63	57.3	15	23 ABB98657	Dextrane-saccharas
36	57	51.8	15	23 ABB98652	Dextrane-saccharas
37	55	50.0	15	23 ABB98656	Dextrane-saccharas
38	51	46.4	855	23 ABB98573	Dextran saccharase
39	48	43.6	479	22 ABB20960	Novel human diagno
40	47	42.7	541	8 AAF70433	Human atrial natrl
41	47	42.7	541	17 AAR95586	Human atrial natrl
42	47	42.7	639	20 AA117089	Artial natriuretic
43	45	40.9	1335	14 AAR42839	Bacillus alkaline
44	44.5	40.5	986	20 AA137617	Urea amidolysase.
45	44	40.0	267	22 ABB24233	Protein which is s
					Novel human diagno

ALIGNMENTS

RESULT 1	AAU98028	standard; Protein; 1375 AA.
XX	AAU98028;	
AC	27-AUG-2002	(first entry)
XX		
DE	S. mutans glucosyltransferase GTFC.	
XX		
KW	Glucosyltransferase; GTFC; transgenic plant; paper sizing;	
KM	coating composition; glucan; starch; latex; thermoplastic molecule;	
KW	amyloplast; vacuole; paper manufacture.	
XX		
OS	Streptococcus mutans.	
XX		
XX	US2002031826-A1.	
XX		
PD	14-MAR-2002.	
XX		
PF	19-DEC-2000; 2000US-0740274.	
XX		
PR	11-DEC-1998; 98US-0210361.	
PR	07-JUN-1995; 95US-0478704.	
PR	07-JUN-1995; 95US-0482711.	
PR	07-JUN-1995; 95US-0485243.	
PR	16-JAN-1998; 98US-0007999.	
PR	16-JAN-1998; 98US-0008172.	
XX	20-JAN-1998; 98US-0009620.	
XX		
PA	(NICH//) NICHOLS S E.	
XX		

[illegible]

```

XX JP2002114709-A.
XX
XX
XX 16-APR-2002.
XX
XX PF 04-OCT-2000; 2000JP-0304889.
XX PR 04-OCT-2000; 2000JP-0304889.
XX PA (UNIV-) UNIV NIPPON.
XX DR WPI; 2002-448101/48.
XX PT Anti-carries agent composed of a monoclonal antibody against an
XX inhibitory enzyme against water insoluble glucan synthetase of glucosyl
XX transferase-B (GTF-B) of Streptococcus mutans -
XX
XX Disclousure; Page 22-25; 28pp; Japanese.
XX
XX The invention relates to a monoclonal antibody against dental caries and
XX CC an anti-carries agent composed of a monoclonal antibody produced by
XX CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
XX CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
XX CC inhibible activity against water insoluble glucan synthetase of glucosyl
XX CC transferase-B. The monoclonal antibody specifically inhibits water
XX CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
XX CC transferase-B and is used in the immunotherapy of dental caries. This
XX CC sequence represents a Streptococcus mutans monoclonal antibody-related
XX CC protein.
XX
SQ Sequence 1375 AA;
Query Match 100.0%; Score 110; DB 23; Length 1375;
Best Local Similarity 100.0%; Pred. No. 1,2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 1 VPSYSFIRAHDSVEVDLIRNTI 22
| | | | | | | | | | | | | | | | | | | | |
DB 578 VPSYSFIRAHDSVEVDLIRNTII 599
RESULT 3
AAU79285
ID AAU79285 standard; Protein; 1017 AA.
XX
XX AC AAU79285;
XX DT 13-ANG-2002 (first entry)
XX DE Streptococcus mutans monoclonal antibody-related protein #2.
XX KM Antibody; dental caries; water insoluble glucan synthetase;
XX KM anti-carries; glucosyl transferase-B; immunotherapy.
XX OS Streptococcus mutans.
XX PN JF2002114709-A.
XX PD 16-APR-2002.
XX PF 04-OCT-2000; 2000JP-0304889.
XX PR 04-OCT-2000; 2000JP-0304889.
XX PA (UNIV-) UNIV NIPPON.
XX DR WPI; 2002-448101/48.
XX PT Anti-carries agent composed of a monoclonal antibody against an
XX inhibitory enzyme against water insoluble glucan synthetase of glucosyl
XX transferase-B (GTF-B) of Streptococcus mutans -
XX
XX Claim 4; Page 17-19; 28pp; Japanese.

```

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX
 SO Sequence 1017 AA;
 Query Match 95.5%; Score 105; DB 23; Length 1017;
 Best Local Similarity 95.5%; Pred. No. 6.2e-09;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 VPSYSFIRAHDSFVQDLIRNII 22
 518 VPSYSFIRAHDSFVQDLIRDII 539
 Db
 RESULT 4
 AAU79284
 ID AAU79284 standard; Protein; 1476 AA.
 AC AAU79284;
 XX
 DT 13-AUG-2002 (first entry)
 DE Streptococcus mutans monoclonal antibody-related protein #1.
 XX
 KM Antibody; dental caries; water insoluble glucan synthetase;
 KM anti-carries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 PN JF2002114709-A.
 PD 16-APR-2002.
 PF 04-OCT-2000; 2000JP-0304889.
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PA (UYN1-) UNIV NIPPON.
 XX
 PS WPI; 2002-448101/48.
 DR
 XX Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans
 XX
 PS Claim 3; Page 13-16; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX
 SO Sequence 1476 AA;
 Query Match 95.5%; Score 105; DB 23; Length 1476;
 Best Local Similarity 95.5%; Pred. No. 9.6e-09;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPSYSFIRAHDSFVQDLIRNII 22
 Db 552 VPSYSFIRAHDSFVQDLIRDII 573
 RESULT 5
 AAU98027
 ID AAU98027 standard; Protein; 1475 AA.
 XX
 AC AAU98027;
 XX
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 DR N-PADB; ABK52938.
 XX
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, M471D, M471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 CC
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPSYSFIRAHDSVEVDLIRNII 22
 DB 552 VPSYSFIRAHDSVEVDLIRNII 573
 RESULT 6
 AAU98030 ID AAU98030 standard; Protein; 1475 AA.
 XX
 XX AAU98030;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant I448V.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 KM
 XX Streptococcus mutans.
 OS
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 448
 FT /note= "Wild-type Ile substituted by Val"
 FT
 XX US2002031826-A1.
 PN
 XX
 XX 14-MAR-2002.
 PD
 XX
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007939.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PR
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PT
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T.
 CC Y168A/Y170R/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPSYSFIRAHDSVEVDLIRNII 22
 DB 552 VPSYSFIRAHDSVEVDLIRNII 573
 RESULT 7
 AAU98031 ID AAU98031 standard; Protein; 1475 AA.
 XX
 XX AAU98031;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 KM
 XX Streptococcus mutans.
 OS
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT
 XX US2002031826-A1.
 PN
 XX
 XX 14-MAR-2002.
 PD
 XX

PF 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSVSEIRHSEVODLLRNTI 22
 DB 552 VPSVSEIRHSEVODLLRNTI 573

RESULT 8
 AAU98033
 ID AAU98033 standard; Protein; 1475 AA.
 XX
 AC AAU98033;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant K1014T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 1014
 FT /note= "wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 DB 552 VPSYSFIRAHDSVQDLIRNII 573

RESULT 9
 AAU98039
 ID AAU98039 standard; Protein; 1475 AA.

AC AAU98039;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant YYY169-171AA.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 169..171 /note= "Wild-type Tyr-Tyr-Tyr substituted by
 FT Ala-Ala-Ala"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions

XX Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 148V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790Q/K1014T,
 CC Y168A/Y170A/Y171A, and K790Q or a GTF D polypeptide having
 CC changes at positions from 1589D, T589E, N471D, N471D/1589D, and
 CC N471D/1589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations and/or vacuole or a maize line
 CC the glucan is produced in the amyloplast and/or vacuole of a plant
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 DB 552 VPSYSFIRAHDSVQDLIRNII 573

RESULT 10
 AAU98040
 ID AAU98040 standard; Protein; 1475 AA.

AC AAU98040;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant K779Q.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"

XX US2002031826-A1.

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.0%; Pred. No. 1e-07; Indels 0; Gaps 0;
 Matches 20; Conservative 1; Mismatches 1;
 1 VPSYSFIRAHDSVQDLIRNII 22
 |||

DB 552 VPSYSFIRAHDSVQDLIRNII 573
 RESULT 11
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.
 XX
 AC AAU98035;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 89.1%; Score 98; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1.5e-07;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 DB 552 VPSYSFIRAHDSVQDLIRNII 573

RESULT 12

AA32925 AAR32925 standard; Protein; 1592 AA.

AC AAR32925;

DT 28-JUN-1993 (first entry)

DE Glucosyltransferase I.

KW GT-1; Streptococcus; dental; caries.

KM Streptococcus sobrinus.

OS JP05023188-A.

PN 02-FEB-1993.

PD 25-JUL-1991; 91JP-0186592.

PR 25-JUL-1991; 91JP-0186592.

XX (FUKU/) FUKUI I.

PA (KATO/) KATO K.

XX WPI; 1993-079449/10.

DR N-PSDB; AAQ37760.

XX DNA sequence glucosyltransferase-I - comprises Streptococcus

PT sobrinus DNA sequence with at least one nucleotide added or

PT deleted

XX Claim 13; Page 15; 29pp; Japanese.

XX The DNA sequence from Streptococcus sobrinus strain 6715 encodes
 CC glucosyltransferase-I (and mutants). The DNA was obtd. by treating
 CC S. sobrinus 6715 with mutanolysin, extracting the chromosomal DNA,
 CC partially digesting with Sau3AI and fractionating on agarose gel.
 CC The 3-5 kbp fragment was ligated into pUC18 and E. coli JM109
 CC transformed with it. A GT-1 expressing clone was isolated and
 CC sequenced. The clone may be used in the development of a drug for
 CC dental caries.

SQ Sequence 1592 AA;

Query Match 89.1%; Score 98; DB 14; Length 1592;
 Best Local Similarity 86.4%; Pred. No. 1.6e-07;
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 DB 548 VPSYSFIRAHDSVQDLIRNII 569

RESULT 13

AAU98032 AAU98032 standard; Protein; 1475 AA.

AC AAU98032;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KW coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.

OS Synthetic.

FN US2002031826-A1.

XX 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 7589D, 7589E, N471D, N471D/7589D, and
 CC N471D/7589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the expression cassette, host cell
 CC promoter, a vector comprising a transgenic plant comprising the
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

KM Glucosyltransferase; GFPB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT US2002031826-A1.
 XX
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 148V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC Properties and impact gloss to the paper during coating step.
 CC The present sequence represents a GFPB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 SQ Sequence 1475 AA;
 Query Match 82.7%; Score 91; DB 23; Length 1475;
 Best Local Similarity 86.4%; Pred. No. 2.4e-06;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VPESYFIRAHDSFVDLRNII 22
 DB 552 VPESYFIRAHDSFVDTLAKII 573
 Search completed: November 13, 2003, 09:38:28
 Job time : 40.3507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 13.6588 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-15
Perfect score: 110
Sequence: 1 VPSYSFIRAHDSFVODLRNII 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	1375	2 JT0345	dextranucrase (EC
2	99	90.0	1475	2 B3135	glTF protein precu
3	98	89.1	1592	2 A38175	glucosyltransferas
4	83	75.5	1290	2 JC5473	dextranucrase (EC
5	83	75.5	1508	2 T31098	probable dextranu
6	76	69.1	1365	2 A41483	glucosyltransferas
7	71	64.5	1449	2 T30857	glucosyltransferas
8	70	63.6	1577	2 T30852	glucosyltransferas
9	70	63.6	1577	2 T30858	glucosyltransferas
10	68	61.8	1431	2 A45866	dextranucrase (EC
11	68	61.8	1518	2 A44811	glucosyltransferas
12	68	61.8	1599	2 S22737	glucosyltransferas
13	50	45.5	51	2 G62455	hypothetical prote
14	48	43.6	540	1 OYHUCR	hypothetical peptid
15	46.5	42.3	414	2 T66120	naturalistic peptid
16	46	41.8	196	2 S48302	hypothetical prote
17	46	41.8	597	2 P82935	MD2 protein - yea
18	46	41.8	631	2 H90599	DNA polymerase III
19	45.5	41.4	767	2 T39715	conserved hypotet
20	45.5	41.4	956	2 H81623	probable transcrip
21	45	40.9	136	2 T12823	conserved hypotet
22	45	40.9	1835	2 S46082	hypothetical prote
23	45	40.9	954	2 T31679	urea carboxylase (
24	44.5	40.5	202	2 E72688	bactitracin synthe
25	44	40.0	282	2 S47795	hypothetical prote
26	44	40.0	327	2 H59094	hypothetical prote
27	44	40.0	327	2 D86474	probable transcrip
28	44	40.0	344	2 JH0511	probable RING zinc
29	44	40.0	344	2 JH0511	inositol 2-dehydro

30	44	40.0	570	2 B97160	fibronectin-bindin
31	44	40.0	587	2 S63033	hypothetical prote
32	44	40.0	734	2 A82047	primosomal replica
33	44	40.0	1068	1 A43322	1-phosphatidylinos
34	44	40.0	1068	1 I38110	1-phosphatidylinos
35	44	40.0	1583	2 F97846	hypothetical prote
36	43	39.1	187	2 H97184	nitroreductase fam
37	43	39.1	214	2 R6BSCA	regulatory protease
38	43	39.1	635	2 D86265	hypothetical prote
39	43	39.1	691	2 S48390	hypothetical prote
40	43	39.1	832	2 C86431	hypothetical prote
41	43	39.1	1260	2 A86323	7518.5 protein - A
42	42	38.2	108	2 S75540	protein P14D16.3 (
43	42	38.2	145	2 G69098	hypothetical prote
44	42	38.2	225	2 S19428	hypothetical prote
45	42	38.2	508	2 T50180	SRD1 protein - yea
					nucleolar protein

ALIGNMENTS

RESULT 1
JT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N/Alternate names: sucrose 6-glucosyltransferase
C/Species: Streptococcus mutans
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C/Accession: JT0345; C31135
R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A/Title: Sequence analysis of the glTF gene from Streptococcus mutans GS-5.
A/Reference number: JT0345; PMID:89137980; PMID:2976010
A/Accession: JT0345
A/Molecule type: DNA
A/Residues: 1-1375 <UED>
A/Experimental source: GS-5
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the glTF gene from Streptococcus mutans.
A/Reference number: A33135; PMID:87308013; PMID:3040685
A/Accession: C31135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-349 <SHI>
A/Cross-references: GS:W17361
C/Genetic: GS
C/Function: glTF
A/Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C/Superfamily: cpl repeat homology
C/Keywords: duplication; glycosyltransferase; hexosyltransferase
F/1-34/Domain: signal sequence #status predicted <SIG>
F/35-1375/Product: glucosyltransferase #status predicted <MAT>
F/1126-1145/Domain: cpl repeat homology <CP1>
F/1253-1272/Domain: cpl repeat homology <CP2>
F/1318-1337/Domain: cpl repeat homology <CP3>

Query Match
Best Local Similarity 100.0%; Score 110; DB 2; Length 1375;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVODLRNII 22
DB 578 VPSYSFIRAHDSFVODLRNII 599

RESULT 2
B3135
glTF protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C/Accession: B3135; A33128
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.

J. Bacteriol. 169, 4263-4270, 1987
 A>Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: B33135
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1475 <SH1>
 A:Cross-references: GB:M17361; NID:g153639; PIDN:AAA8586.1; PID:g153640
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 Submitted to the Protein Sequence Database, September 1990
 A:Reference number: A33128
 A:Accession: A33128
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-171,173-641, 'N', 643-1475 <SH2>
 A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 90.0%; Score 99; DB 2; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 3.2e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSTRANDSEVQDILIRNII 22
 DB 552 VPSYSTRANDSEVQDILIRNII 573

RESULT 3

A38175
 Glucosyltransferase precursor - Streptococcus sobrinus
 C:Species: Streptococcus sobrinus
 C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
 C:Accession: A38175
 J:Abou, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
 J. Bacteriol. 173, 989-996, 1991
 A>Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
 A:Reference number: A38175; MUID:91123227; PMID:1704006
 A:Accession: A38175
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1592 <ABO>
 A:Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g217033
 C:Superfamily: cpl repeat homology
 F:1093-1112/Domain: cpl repeat homology <CP1>
 F:1222-1241/Domain: cpl repeat homology <CP2>
 F:1287-1306/Domain: cpl repeat homology <CP3>
 F:1330-1351/Domain: cpl repeat homology <CP4>
 F:1352-1371/Domain: cpl repeat homology <CP5>
 F:1402-1420/Domain: cpl repeat homology <CP6>
 F:1465-1484/Domain: cpl repeat homology <CP7>
 F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 89.1%; Score 98; DB 2; Length 1592;
 Best Local Similarity 86.4%; Pred. No. 5.1e-07;
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSTRANDSEVQDILIRNII 22
 DB 548 VPSYSTRANDSEVQDILIRNII 569

RESULT 4

JC5473
 dextranucrase (EC 2.4.1.5) - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C:Accession: JC5473
 R:Monchois, V.; Willemot, R.M.; Renaud-Simeon, M.; Croux, C.; Monsan, P.

Gene 182, 23-32, 1996
 A>Title: Cloning and sequencing of a gene coding for a novel dextranucrase from Leucono
 A:Reference number: JC5473; MUID:97136686; PMID:8982063
 A:Accession: JC5473
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1290 <MON>
 A:Cross-references: GB:U98181
 C:Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose ont
 C:Genetics:
 A:Gene: dxtA
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:78-870/Domain: catalytic #status predicted <CAT>
 F:922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 75.5%; Score 83; DB 2; Length 1290;
 Best Local Similarity 76.2%; Pred. No. 9.8e-05;
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PYSYSTRANDSEVQDILIRNII 22
 DB 388 PYSYSTRANDSEVQDILIRNII 408

RESULT 5
 T31098
 Probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T31098
 R:Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
 FEBS Microbiol. Lett. 159, 307-315, 1998
 A>Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A:Reference number: Z20981; MUID:98164374; PMID:9503626
 A:Accession: T31098
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1508 <MON>
 A:Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AA895453.1
 A:Experimental source: strain NRRL B-1299
 C:Genetics:
 A:Gene: dxtB
 C:Function: produces dextran composed only of alpha(1-6) glucosidic bonds
 A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 75.5%; Score 83; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.00012;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSTRANDSEVQDILIRNII 22
 DB 634 IPNYSFVRAHDSVQTVIAQII 655

RESULT 6

A41483
 glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus
 C:Species: Streptococcus sobrinus
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C:Accession: A41483
 R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
 Infect. Immun. 58, 2452-2458, 1990
 A>Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltrans
 A:Reference number: A41483; MUID:90316665; PMID:2142479
 A:Accession: A41483
 A:Molecule type: DNA
 A:Residues: 1-1365 <GIL>
 A:Cross-references: GB:M30943; NID:g153652; PIDN:AAA26898.1; PID:g153653
 C:Genetics:
 A:Gene: gtfS
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 69.1%; Score 76; DB 2; Length 1365;
Best Local Similarity 72.7%; Pred. No. 0.0014;
Matches 16; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VPSYFIRAHSEVQDLIRNII 22
||:|||||||:|
Db 537 VNYVFIKRAHSEVQTRIAKII 558

RESULT 7

T30857
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
I: Infect. Immun. 63, 609-621, 1995
A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for p
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30857
A:Status: preliminary; translated from GB/EMBL/DBJ
C:Keywords: glycosyltransferase; hexosyltransferase
A:Residues: 1-1449 <SIM>
A:Molecule type: DNA
A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AA041412.1
C:Genetics:
A:Gene: gltF

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.0093;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
||:|||||||:|
Db 609 NYAFVRAHSEVQSIIGQII 628

RESULT 8

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaffe, R.I.
I: submitted to the EMBL Data Library, February 1998
A:Description: Streptococcus salivarius VI477 gltF.
A:Reference number: Z20854
A:Accession: T30552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <JAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AA05156.1
C:Genetics:
A:Gene: gltF

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.0093;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
||:|||||||:|
Db 609 NYAFVRAHSEVQSIIGQII 628

RESULT 9

T30858
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
I: Infect. Immun. 63, 609-621, 1995
A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for p

A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>
A:Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIDN:AA041413.1
C:Genetics:
A:Gene: gltF

Query Match 63.6%; Score 70; DB 2; Length 1577;
Best Local Similarity 65.0%; Pred. No. 0.015;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
||:|||||||:|
Db 661 NYIFVRAHSEVQAVLANII 680

RESULT 10

A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A45866
R:Honda, O.; Kato, C.; Kuramitsu, H.K.
J. Gen. Microbiol. 136, 2099-2105, 1990
A:Title: Nucleotide sequence of the Streptococcus mutans gltD gene encoding the glucosyl
A:Reference number: A45866; MUID:91100958; PMID:2148600
A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>

A:Cross-references: GB:M29296
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1257-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>
F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 61.8%; Score 68; DB 2; Length 1431;
Best Local Similarity 70.0%; Pred. No. 0.028;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
||:|||||||:|
Db 576 NYIFVRAHSEVQTVIAKII 595

RESULT 11

A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811
A:Molecule type: DNA
A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAA77900.1; PID:g47527
A:Note: sequence extracted from NCBI Backbone (NCBIN:81050, NCBIP:81052)
C:Genetics:
A:Gene: gltF
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 61.8%; Score 68; DB 2; Length 1518;
Best Local Similarity 65.0%; Pred. No. 0.03;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRADSEVODLIRNII 22
DB 604 NYFIRAHNNQDIIAEII 623

RESULT 12
S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C/Species: Streptococcus salivarius
C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C/Accession: S22737; S28810; B44811; S22727
R/Jacques, N.

Submitted to the EMBL Data Library, March 1992
A/Reference number: S22726

A/Accession: S22737

A/Molecule type: DNA
A/Residues: 1-1599 <JAC>

A/Cross-references: EMBL:Z11872; NID:G47530; PIDD:CAA77898.1; PID:G47531

A/Experimental source: ATCC 25975

R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A/Title: Molecular characterization of a cluster of at least two glucosyltransferase genes

A/Reference number: A44811; MUID:Z2148377; PMID:1838361

A/Accession: S28810

A/Molecule type: DNA
A/Residues: 1-51 <GIF>

A/Cross-references: EMBL:Z11873

C/Genetics:

A/Gene: gtfK
C/Superfamily: cpl repeat homology

C/Keywords: glycosyltransferase; hexosyltransferase
F/1456-1475/Domain: cpl repeat homology <CPR>

Query Match 61.8%; Score 68; DB 2; Length 1599;
Best Local Similarity 65.0%; Pred. No. 0.031;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRADSEVODLIRNII 22
DB 574 TYLFRAHDSVQVIADII 593

RESULT 13
G82455
hypothetical protein VCA0471 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: G82455

R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodeson, R.J.;

Charadon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Driegol, I.; Sellers, F.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: G82455

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-51 <HRI>

A/Cross-references: GB:AE004379; GB:AE003853; NID:G9657865; PIDD:AAF96375.1; GSPDB:GN001

A/Experimental source: serogroup O1, strain N16961, biotype El Tor

C/Genetics:

A/Gene: VCA0471

A/Map position: 2

Query Match 45.5%; Score 50; DB 2; Length 51;
Best Local Similarity 54.5%; Pred. No. 0.43;
Matches 12; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

QY 1 VPSYFIRADSEVODLIRNII 22
DB 15 VPAISAIR--NSRIRDLKXRII 34

RESULT 14
OYHUCR
natriuretic peptide receptor C precursor - human

N/Alternate names: atrial natriuretic factor clearance receptor; natriuretic peptide cle

C/Species: Homo sapiens (man)

C/Date: 09-Nov-1990 #sequence_revision 05-May-1995 #text_change 22-Jun-1999

C/Accession: S10150; A35896

R/Lowe, D.G.; Camarato, T.R.; Goeddel, D.V.

Nucleic Acids Res. 18, 3412, 1990

A/Title: cDNA sequence of the human atrial natriuretic peptide clearance receptor.

A/Reference number: S10150; MUID:90287735; PMID:2162522

A/Accession: S10150

A/Molecule type: mRNA

A/Residues: 1-540 <LOW>

A/Cross-references: EMBL:X52282; NID:G28705; PIDD:CAA36523.1; PID:G28706

A/Note: alternative splice form C6

R/Porter, J.G.; Arsten, A.; Fuller, F.; Miller, J.A.; Gregory, L.C.; Lewicki, J.A.

Biochem. Biophys. Res. Commun. 171, 796-803, 1990

A/Title: Isolation and functional expression of the human atrial natriuretic peptide cle

A/Reference number: A35896; MUID:90386656; PMID:2169733

A/Accession: A35896

A/Molecule type: mRNA

A/Residues: 1-475, 'SG', 477-540 <POR>

A/Cross-references: GB:M59305; NID:G178651; PIDD:AA51734.1; PID:G178652

A/Note: alternative splice form C5

R/Shuler, J.T.; O'Connell, K.L.; Garcia, C.; Wong, S.; Engel, A.M.; Garbers, D.L.; Lowe,

Biochemistry 33, 11372-11381, 1994

A/Title: The disulfide linkages and glycosylation sites of the human natriuretic peptide

A/Reference number: A55870; MUID:95244450; PMID:7727388

A/Contents: annotation; propeptide, disulfide bonds, carbohydrate binding sites

C/Genetics:

A/Gene: GDB:NR3; NRPC; ANRPC

A/Cross-references: GDB:125201; OMIM:108962

A/Map position: 5p14-5p13

C/Superfamily: natriuretic peptide receptor C; natriuretic peptide-binding domain homolo

C/Keywords: alternative splicing; glycoprotein; homodimer; hormone receptor; transmembr

F/1-23/Domain: signal sequence #status predicted <SIG>

F/24-45/Domain: propeptide #status experimental <PRO>

F/46-540/Product: natriuretic peptide receptor C #status predicted <EXT>

F/46-479/Domain: extracellular #status predicted <TMM>

F/131-462/Domain: transmembrane #status predicted <TMM>

F/483-503/Domain: transmembrane #status predicted <TMM>

F/108-136, 213-261/Disulfide bonds: #status experimental

F/394/Binding site: carbohydrate (Asn) (covariant) (partial) #status experimental

F/473-476/Disulfide bonds: (partial) (in splice form C6) #status experimental

Query Match 43.6%; Score 48; DB 1; Length 540;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 YSFIRADSEVODLIRNII 21
DB 224 YSFDETKDLDLIDYIRNI 251

RESULT 15
T06120
hypothetical protein F23E12.70 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999

C/Accession: T06120

R/Revan, M.; Hilbert, H.; Braun, M.; Holzer, B.; Brandt, A.; Duesterhoeft, A.; Hobeisel,

submitted to the Protein Sequence Database, April 1999

A/Reference number: Z15485

A/Accession: T06120

A/Molecule type: DNA
A/Residues: 1-414 <BEV>

A:Cross-references: EMBL:AL022604; GSPDB:GN00062; ATSP:F23E12.70
 A:Experimental source: cultivar Columbia; BAC clone F23E12
 C:Genetics:
 A:Gene: ATSP:F23E12.70
 A:Map position: 4
 A:introns: 83/1; 109/3; 180/1; 199/3; 215/3; 259/3; 284/3; 317/3; 359/3
 Query Match 42.3%; Score 46.5; DB 2; Length 414;
 Best Local Similarity 55.0%; Pred. No. 18;
 Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
 Qy 2 PSYFIRAHDSVQDILIRNI 21
 ||:|||||:
 Db 337 PSF-IIHAHSEVSSISYNI 355
 Search completed: November 13, 2003, 09:50:28
 Job time : 13.6588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 / Search time 7.92417 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-15

Perfect score: 110
Sequence: 1 VPSYSFIRAHSEVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	110	100.0	1455	1	GFPC_STRMU
2	105	95.5	1476	1	GFPC_STRMU
3	100	90.9	1597	1	GFPC_STRMU
4	98	89.1	1592	1	GFPC_STRMU
5	96	87.1	1365	1	GFPC_STRMU
6	68	61.8	1462	1	GFPC_STRMU
7	48	43.6	541	1	ANPC_HUMAN
8	46	41.8	1835	1	MAD2_YEAST
9	45	40.9	1835	1	MAD2_YEAST
10	45	40.9	6359	1	BACC_BACCI
11	44	40.0	344	1	YTAU_ECOLI
12	44	40.0	344	1	YTAU_ECOLI
13	44	40.0	1068	1	P11A_BOVIN
14	44	40.0	1068	1	P11A_BOVIN
15	44	40.0	1068	1	P11A_BOVIN
16	44	40.0	1068	1	P11A_BOVIN
17	44	40.0	1068	1	P11A_BOVIN
18	43	39.1	397	1	LEI1_METRA
19	43	39.1	626	1	MP44_FOWPY
20	43	39.1	691	1	YIO4_YEAST
21	43	39.1	2541	1	TMN1_HUMAN
22	43	39.1	225	1	SRP1_YEAST
23	42	38.2	584	1	SYD_BUTAP
24	42	38.2	932	1	HIR1_SCHPO
25	42	38.2	4960	1	MDN1_YEAST
26	42	38.2	4960	1	MDN1_YEAST
27	41.5	37.7	713	1	DCU2_ECOLI
28	41.5	37.7	713	1	DCU2_ECOLI
29	41.5	37.7	713	1	DCU2_ECOLI
30	41	37.3	183	1	YS4L_CABEL
31	41	37.3	242	1	SGAE_MYCEN
32	41	37.3	354	1	MBHS_AZOCH
33	41	37.3	537	1	ANPC_BOVIN

34	41	37.3	659	1	RNB_HABIN	P44440 haemophilus
35	41	37.3	826	1	TLR4_PAPAN	Q95222 papio anubi
36	41	37.3	875	1	ENV_BIV06	P19556 bovine immu
37	41	37.3	904	1	ENV_BIV27	P19557 bovine immu
38	41	37.3	1006	1	KL25_TOBAC	O23826 nicotiana t
39	41	37.3	2542	1	KL25_TOBAC	Q94966 homo sapien
40	40.5	36.8	183	1	RET2_BOVIN	P09517 barley yell
41	40.5	36.8	201	1	RET2_BOVIN	P18902 sus scrofa
42	40.5	36.8	96	1	IBB2_PEA	Q41065 pisum sativ
43	40	36.4	114	1	IBB2_PEA	Q41066 pisum sativ
44	40	36.4	205	1	ADEN_ADEG8	Q9gm72 avian adeno
45	40	36.4	205	1	ADEN_ADEG8	

ALIGNMENTS

RESULT 1
GFPC_STRMU STANDARD; PRT: 1455 AA;
ID GFPC_STRMU
AC P13470; 069382; 069385; 069391; 069397; P05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
GN (Dextranucrase) (Sucrose 6-glucosyltransferase).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1309;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5";
RL Gene 69:101-109(1988).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype B, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=U04159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian X.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Dia H., Lin S., Qian X.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perlecci J.J.;
RT "Genome sequence of Streptococcus mutans U04159, a cariogenic dental
RT pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
[4]
RN
RP SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans";
RL J. Bacteriol. 169:4263-4270(1987).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)1(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)1(N)+1.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.

```

FT FT VARIANT 1398 1398 D -> N (IN STRAIN MT42239) .
FT FT VARIANT 1424 1424 V -> I (IN STRAINS MT42239 AND MT8148) .
FT FT VARIANT 1439 1439 S -> P (IN STRAIN MT8148) .
FT FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148) .
FT FT CONFLICT 1337 1455 ORLTFKNGVQAKELITERRGRILKYIDPNSGNEVRNRYR
TT TT QSLVFKNGVQAKELITGMHVEGRYVDENGYRIASHD
TT TT TSGNVNYFGNDGYALIGMTHVGRVDFENGYRIASHD
TT TT ORNMWDYRDYRDPGRSSAVFRSRNGFEEDNPRF ->
TT TT HASLIISLMVPTLRSSLSGSVKVVSNTMLIIPEMKREIVYM
TT TT (IN REF.1).
SQ SEQUENCE 1455 AA; 162965 NM; 3CB455A99AAFEBC86 CRC64;
Query Match 100.0%; Score 110; DB 1; Length 1455;
Best Local Similarity 100.0%; Prod. No. 2-6e-09;
Matches 22; Conservatly 0; Mismatches 0; Indels 0; Gaps 0
CY 1 VPSSYFIARHSEVODLIRNI 22
DB 578 VPSSYFIARHSEVODLIRNI 599

RESULT 2
GTFB_STRMU STANDARD: PRT; 1476 AA.
ID GTFB_STRMU STANDARD: PRT; 1476 AA.
AC P089387; O69381; O69384; O69390; O69396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE GLUCOSYLTRANSFERASE-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase) .
GN GTFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RL "Sequence analysis of the gtfB gene from Streptococcus mutans.";
J. Bacteriol. 169:4263-4270(1987).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans."
RT FEWS Microbiol. Lett. 161:331-336(1998).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mcshan W.M., McLoughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar C., Lai H., White J., Roe B.A., Ferretti J.J.;
RA "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen."
RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002) .
RL - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
RL TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC - CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl}(N+1) .
CC - SUBCELLULAR LOCATION: Secreted.
CC - DISEASE: DENTAL CARIES.
CC MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE) . GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.

```

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M17361; AAA8588.1; -

DR EMBL; D88651; BAA26101.1; -

DR EMBL; D88654; BAA26105.1; -

DR EMBL; D88657; BAA26109.1; -

DR EMBL; D88660; BAA26113.1; -

DR EMBL; D89977; BAA26119.1; -

DR EMBL; AE014940; AAN58705.1; -

DR InterPro; IPR002479; CW_binding_70.

DR InterPro; IPR003318; Glyco_hydro_70.

DR Pfam; PF01473; CW_binding_1; 1.

DR Pfam; PF02324; Glyco_hydro_70; 1.

DR transferase; Glycosyltransferase; Signal; Repeat; Dental caries; KW

KW Complete proteome.

FT SIGNL 1 34 POTENTIAL.

FT CHAIN 35 1476 GLUCOSYLTRANSFERASE-I.

FT DOMAIN 35 1051 CATALYTIC (APPROXIMATE).

FT DOMAIN 1097 1476 GLUCAN-BINDING (APPROXIMATE).

FT REPEAT 1097 1130 A REPEAT.

FT REPEAT 1161 1470 5 X TANDEM REPEATS.

FT REPEAT 1161 1210 1.

FT REPEAT 1225 1275 2.

FT REPEAT 1290 1340 3.

FT REPEAT 1355 1405 4.

FT REPEAT 1420 1470 5.

FT REPEAT 1420 1470 S -> T (IN STRAIN MT4239).

FT REPEAT 1420 1470 S -> I (IN STRAIN GS-5).

FT REPEAT 1420 1470 V -> A (IN STRAINS GS-5, MT4245, MT4251, MT4467 AND MT8148).

FT REPEAT 1420 1470 Q -> P (IN STRAIN MT4251).

FT REPEAT 1420 1470 I -> S (IN STRAINS GS-5, MT4245, MT4251, MT4467 AND MT8148).

FT REPEAT 1420 1470 S -> F (IN STRAIN MT4251).

FT REPEAT 1420 1470 K -> N (IN STRAIN MT4251).

FT REPEAT 1420 1470 S -> D (IN STRAINS GS-5, MT4467 AND MT8148).

FT REPEAT 1420 1470 N -> R (IN STRAIN MT4239).

FT REPEAT 1420 1470 I -> T (IN STRAIN MT4239).

FT REPEAT 1420 1470 K -> R (IN STRAIN MT8148).

FT REPEAT 1420 1470 F -> Y (IN STRAIN MT8148).

FT REPEAT 1420 1470 T -> I (IN STRAIN MT8148).

FT REPEAT 1420 1470 A -> V (IN STRAIN MT8148).

FT REPEAT 1420 1470 F -> L (IN STRAIN MT8148).

FT REPEAT 1420 1470 FGGPVE -> YGTPVA (IN STRAINS GS-5, MT4239 AND MT4467).

FT REPEAT 1420 1470 SV -> NT (IN STRAINS GS-5, MT4239 AND MT4467).

FT REPEAT 1420 1470 ADS -> VDG (IN STRAINS GS-5, MT4239 AND MT4467).

FT REPEAT 1420 1470 A -> T (IN STRAIN MT4239).

FT REPEAT 1420 1470 S -> N (IN STRAIN MT4239).

FT REPEAT 1420 1470 H -> Y (IN STRAIN MT4251).

FT REPEAT 1420 1470 E -> K (IN STRAIN MT8148).

FT REPEAT 1420 1470 Y -> C (IN STRAIN MT8148).

FT REPEAT 1420 1470 A -> P (IN STRAIN MT4239).

FT REPEAT 1420 1470 K -> P (IN STRAIN MT8148).

FT REPEAT 1420 1470 R -> H (IN STRAINS GS-5 AND MT4467).

FT REPEAT 1420 1470 Y -> H (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).

FT REPEAT 1420 1470 S -> G (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).

FT REPEAT 1420 1470 H -> Y (IN STRAINS GS-5 AND MT4467).

FT VARIANT 1394 1394 Y -> H (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).

FT VARIANT 1402 1402 S -> G (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).

FT VARIANT 1459 1459 Y -> H (IN STRAIN MT4467).

FT CONFLICT 570 570 R -> A (IN REF. 1).

FT CONFLICT 800 817 ADGCVASAPAPSTDGK -> LTKMFLRLRPQDWA (IN REF. 1).

FT CONFLICT 1310 1310 H -> L (IN REF. 1).

FT CONFLICT 1476 AA; 165846 MM; 90CE09F731B4C6CF CRC64; SQ SEQUENCE

Query Match 95.5%; Score 105; DB 1; Length 1476; Best Local Similarity 95.5%; Pred. No. 1.6e-08; Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYFIRHSEVODLRITII 22

DB 552 VPSYFIRHSEVODLRITII 573

RESULT 3

GTFL_STRDO STANDARD; PRT; 1597 AA.

ID GTFL_STRDO

AC P11001; (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)

DE (Sucrose 6-glucosyltransferase).

GN GTFI.

OS Streptococcus downei (Streptococcus sobrinus).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1317;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MFE28;

RX MEDLINE=87308014; PubMed=3040686;

RA Ferretti J.D., Gilpin M.L., Russell R.B.;

RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus sobrinus MFE28";

RL J. Bacteriol. 169:4271-4278 (1987).

CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-fructose + [(1,6)-alpha-D-glucosyl] (N+1).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M17391; AAC63063.1; -

DR InterPro; IPR002479; CW_binding_70.

DR InterPro; IPR003318; Glyco_hydro_70.

DR Pfam; PF01473; CW_binding_1; 16.

DR Pfam; PF02324; Glyco_hydro_70; 1.

DR transferase; Glycosyltransferase; Signal; Repeat; Dental caries. KW

FT SIGNAL 1 38 POTENTIAL.

```

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1099 1597 TRANSFERASE; 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1132 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MM; B9B6A200868798E CRC64;

Query Match 90.3%; Score 100; DB 1; Length 1597;
Best Local Similarity 90.3%; Pred. No. 1,1e-07;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSEFIRAHDSVODLRNII 22
Db 554 VPSYSEFIRAHDSVODLRDII 575

RESULT 4
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
OK NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Aho H., Matsunura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES BOTH
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BA014241.1;
CC InterPro; IPR002479; CW_binding.

```

```

DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 13.
DR Pfam; PF0324; Glyco_hydro_70; 1.
DR TRANSFERASE; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MM; BC0A6DD079351ECF CRC64;

Query Match 89.1%; Score 98; DB 1; Length 1592;
Best Local Similarity 86.4%; Pred. No. 2,3e-07;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSEFIRAHDSVODLRNII 22
Db 548 VPSYSEFIRAHDSVODLRDII 569

RESULT 5
GTF5_STRDO STANDARD; PRT; 1365 AA.
ID GTF5_STRDO
AC P29336;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
OK NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MPE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell R.R., Ferretti J.J.;
RT "Analysis of the Streptococcus downei gtfS gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans.";
RL Infect. Immun. 58:2452-2458(1990).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30943; AAA26898.1;

```

DR InterPro: IPR002479; CW binding.
DR InterPro: IPR003118; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 8.
DR Pfam: PF03324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36
FT CHAIN 37 1365 GLUCOSYLTRANSFERASE-S.
FT REPEAT 157 177 CELL WALL BINDING 1.
FT REPEAT 178 197 CELL WALL BINDING 2.
FT REPEAT 198 1061 CATALYTIC (APPROXIMATE).
FT DOMAIN 1062 1082
FT REPEAT 1083 1102 CELL WALL BINDING 3.
FT REPEAT 1150 1169 CELL WALL BINDING 4.
FT REPEAT 1170 1190 CELL WALL BINDING 5.
FT REPEAT 1225 1243 CELL WALL BINDING 6.
FT REPEAT 1289 1308 CELL WALL BINDING 7.
FT REPEAT 1309 1328 CELL WALL BINDING 8.
FT REPEAT 1331 1352 CELL WALL BINDING 9.
FT REPEAT 1365 AA; 151590 MW; 1672965A2B8C476 CRC64;
SQ SEQUENCE

Query Match 69.1%; Score 76; DB 1; Length 1365;
Best Local Similarity 72.7%; Pred. No. 0.0006; Indels 0; Gaps 0;
Matches 16; Conservative 1; Mismatches 5;

QY 1 VPSYFIRAHDSVQDLIRNII 22
DB 537 VPMYVIRAHDSVQRIAKII 558

RESULT 6
GTFD_STRMU STANDARD; PRT; 1462 AA.
ID GTFD_STRMU 069383; 069386; 069389; 069392; 069398;
AC P49331; 069383; 069386; 069389; 069392; 069398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme."
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype B, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES BOTH
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@db-sib.ch).
CC
CC EMBL: M29296; AAA26895.1; -
CC EMBL: D88653; BAA26103.1; -
CC EMBL: D88656; BAA26107.1; -
CC EMBL: D88659; BAA26111.1; -
CC EMBL: D88652; BAA26115.1; -
CC EMBL: D89979; BAA26121.1; -
CC EMBL: AE014932; AAN58619.1; -
CC InterPro: IPR002479; CW binding.
CC InterPro: IPR003118; Glyco_hydro_70.
CC Pfam: PF03324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW Complete proteome.
FT CHAIN 1 ?
FT DOMAIN 1232 1462
FT REPEAT 1232 1295
FT REPEAT 1296 1359
FT REPEAT 1360 1423
FT REPEAT 1360 1423
FT VARIANT 10
FT VARIANT 19
FT VARIANT 19
FT VARIANT 58
FT VARIANT 68
FT VARIANT 81
FT VARIANT 113
FT VARIANT 122
FT VARIANT 132
FT VARIANT 135
FT VARIANT 137
FT VARIANT 202
FT VARIANT 255
FT VARIANT 275
FT VARIANT 288
FT VARIANT 301
FT VARIANT 313
FT VARIANT 317
FT VARIANT 328
FT VARIANT 350
FT VARIANT 628
FT VARIANT 688
FT VARIANT 726
FT VARIANT 730
FT VARIANT 762
FT VARIANT 762
POTENTIAL.
GLUCOSYLTRANSFERASE-S
3 X 63 AA APPROXIMATE TANDEM REPEATS.
1.
2.
3.
Y -> H (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
I -> V (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
K -> E (IN STRAIN MT4467).
A -> S (IN STRAINS MT4239 AND MT4245).
A -> T (IN STRAINS MT4251 AND MT8148).
T -> I (IN STRAINS MT4239, MT4245 AND
MT8148).
A -> V (IN STRAINS GS-5 AND MT4467).
A -> S (IN STRAINS GS-5 AND MT4467).
A -> T (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
V -> L (IN STRAIN MT4239).
D -> N (IN STRAIN MT8148).
E -> D (IN STRAINS MT4239, MT4245 AND
MT4251).
D -> N (IN STRAINS MT4239, MT4245 AND
MT4251).
Q -> H (IN STRAIN MT4245).
D -> N (IN STRAINS MT4239 AND MT4251).
E -> K (IN STRAIN MT4239).
V -> F (IN STRAIN MT4239).
F -> L (IN STRAINS MT4239, MT4251 AND
MT4467).
KKKTTQ -> EEEVTL (IN STRAIN MT4251).
A -> S (IN STRAIN MT4239).
IDOGSEA -> ADKGNDS (IN STRAIN MT4251).
IDOGS -> ADKGN (IN STRAINS MT4239 AND
MT4245).
T -> A (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).

```

FT  VARIANT  964  964  D -> Y (IN STRAIN MT4251).
FT  VARIANT  1019 1019  E -> K (IN STRAIN MT4251).
FT  VARIANT  1059 1060  LG -> IR (IN STRAIN MT4251).
FT  VARIANT  1060 1060  G -> R (IN STRAIN MT4245).
FT  VARIANT  1080 1080  G -> R (IN STRAIN MT4239).
FT  VARIANT  1142 1142  H -> Q (IN STRAIN GS-5).
FT  VARIANT  1198 1198  S -> N (IN STRAIN MT4239).
FT  VARIANT  1220 1220  Y -> C (IN STRAIN MT4251 AND MT4467).
FT  VARIANT  1280 1280  Q -> L (IN STRAIN MT4467).
FT  VARIANT  1282 1282  Q -> P (IN STRAIN MT4245).
FT  VARIANT  1290 1290  K -> T (IN STRAIN MT4245).
FT  VARIANT  1311 1311  N -> D (IN STRAIN MT4245).
FT  VARIANT  1403 1403  G -> D (IN STRAIN GS-5 AND MT4467).
FT  VARIANT  1425 1425  G -> R (IN STRAIN GS-5).
FT  VARIANT  1449 1449  R -> K (IN STRAIN MT4467).
FT  CONFLICT 1428 1462  R (IN REF. 1).
SQ  SEQUENCE 1462 AA; 16387 MW; CE4A279C4D708645 CRC64;
Query Match 61.8%; Score 68; DB 1; Length 1462;
Best Local Similarity 70.0%; Pred. No. 0.012; 4; Indels 0; Gaps 0;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 3 YSFIRAHSEVQDLIRNI 22
Db 576 NYFIRAHSEVQTVIAKII 595

RESULT 7
ANPC_HUMAN STANDARD; PRT; 541 AA.
ID ANPC_HUMAN
AC P17342;
DT 01-AUG-1990 (Rel. 15, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Atrial natriuretic peptide clearance precursor (ANP-C)
DE (ANP-C) (Atrial natriuretic peptide C-type receptor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Kidney;
RX MEDLINE=90287735; PubMed=2162522;
RA Lowe D.G., Camerato T.R., Goeddel D.V.;
RT "cDNA sequence of the human atrial natriuretic peptide clearance
RT receptor."
RL Nucleic Acids Res. 18:3412-3412(1990).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP MEDLINE=90386656; PubMed=2169733;
RA Porter J.G., Arsten A., Fuller F., Miller J.A., Gregory L.C.,
RA Lewicki J.A.;
RT "Isolation and functional expression of the human atrial natriuretic
RT peptide clearance receptor cDNA."
RL Biochem. Biophys. Res. Commun. 171:796-803(1990).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Lens epithelium;
RA Rae U.L., Shepard A.R.;
RT "Human lens epithelial mRNA for atrial natriuretic peptide clearance
RT receptor."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DE FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. DOES NOT HAVE
CC -1- FUNCTION: CYCLASE ACTIVITY.
CC -1- SUBUNIT: Homodimer; disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: Type I membrane protein.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P17342-1; Sequence=Displayed;

```

```

CC Name=2;
CC IsoId=P17342-2; Sequence=VSP 001812;
CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC -1- SIMILARITY: TO ANP-A AND ANP-B RECEPTORS IN THEIR EXTRACELLULAR
CC AND TRANSMEMBRANE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52282; CAA36523.1; -
CC EMBL: M59305; AAA51734.1; -
CC EMBL: AF025998; AAB88801.1; -
CC PIR: S10150; OYHUCR.
CC PDB: 1JDN; 05-SEP-01.
CC Genew; HGNC:7945; NPR3.
CC MIM: 108962; -
CC GO: GO:0004872; F:receptor activity; TAS.
CC GO: GO:0001501; P:skeletal development; TAS.
CC InterPro: IPR001828; ANP_receptor.
CC InterPro: IPR001170; Ntpep_receptor.
CC Pfam; PF01094; ANP_receptor; 1.
CC PRINTS; PR00255; NATPEPTIDER.
CC PROSITE; PS00458; ANP_RECEPTORS; 1.
CC Receptor; Glycoprotein; Transmembrane; Signal; Alternative splicing;
CC 3D-structure.
CC FT SIGNAL 1 26
CC FT CHAIN 27 541
CC FT DOMAIN 27 481
CC FT TRANSMEM 482 504
CC FT DISULFID 108 136
CC FT DISULFID 213 261
CC FT DISULFID 473 473
CC FT CARBOHYD 86 86
CC FT CARBOHYD 293 293
CC FT CARBOHYD 394 394
CC FT VARSPLIC 476 477
CC SQ SEQUENCE 541 AA; 59807 MW; 8A66415F7F7D62B7 CRC64;
Query Match 43.6%; Score 48; DB 1; Length 541;
Best Local Similarity 44.4%; Pred. No. 6.2; 5; Indels 0; Gaps 0;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 4 YSFIRAHSEVQDLIRNI 21
Db 234 YSFDETKDLRLDIVANI 251

RESULT 8
MAD2_YEAST STANDARD; PRT; 196 AA.
ID MAD2_YEAST
AC P40958;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mitotic spindle checkpoint component MAD2 (Mitotic MAD2 protein).
DE MAD2 OR YJL030W OR J1256.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.

```

RA Li R., Havel C., Watson J.A., Murray A.M.;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RA Pohl T.M., Aljinovic G.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN
 RP
 RE CHARACTERIZATION.
 RC MEDLINE=9130300; PubMed=1651172;
 RA Li R., Murray A.M.;
 RL "Feedback control of mitosis in budding yeast."
 RT Cell 66:519-531(1991).
 RN
 RP INTERACTIONS.
 RC MEDLINE=98128031; PubMed=9461437;
 RA Hwang L.H., Lau L.F., Smith D.L., Mistro C.A., Hardwick K.G.,
 RA Hwang E.S., Amon A., Murray A.M.;
 RL "Budding yeast Cdc20: a target of the spindle checkpoint."
 RT Science 279:1041-1044(1998).
 CC
 CC -1- FUNCTION: FEEDBACK CONTROL THAT PREVENTS CELLS WITH INCOMPLETELY
 CC ASSEMBLED SPINDLES FROM LEAVING MITOSIS.
 CC -1- SUBUNIT: THE SPINDLE CHECKPOINT COMPLEX IS COMPOSED OF MAD1, MAD2
 CC AND MAD3. IT INTERACTS WITH CDC20.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE MAD2 FAMILY.
 CC -1- SIMILARITY: Contains 1 HORMA domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U14132; AAA21385.1;
 DR EMBL; Z49305; CAA89321.1;
 DR PIR; S48302; S48302.
 DR HSSP; Q13257; IDUJ.
 DR SGD; S0003567; MAD2.
 DR GO; GO:0005699; C:kinetochore; IDA.
 DR GO; GO:0005643; C:nuclear pore; IDA.
 DR GO; GO:0007094; P:mitotic spindle checkpoint; IGI.
 DR InterPro; IPR003511; DNBbind_HORMA.
 DR Pfam; PF02301; HORMA; 1.
 DR PROSITE; PS00815; HORMA; 1.
 DR Cell cycle; Mitosis; Nuclear protein.
 DR PROSITE; PS00815; HORMA.
 FT DOMAIN 8 192
 SQ SEQUENCE 196 AA; 22284 MW; EFE59916C5720644 CRC64;
 QY 5 SFRHSDSVQDLINII 22
 DB 46 TLKTHDELKQYIKIL 63
 Query Match 41.8%; Score 46; DB 1; Length 196;
 Best local similarity 38.9%; Pred. No. 4.3;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92199240; PubMed=1602034;
 RA Genhauffe F.S., Cooper T.G.;
 RL "The urea amidolyase (DUR1,2) gene of *Saccharomyces cerevisiae*."
 RT DNA Seq. 2:19-32(1991).
 RN
 RP SEQUENCE OF 1-893 FROM N.A.
 RC STRAIN=S288c;
 RA Rieger M.;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 873-1835 FROM N.A.
 RC STRAIN=S288c;
 RA Reidmann H., Mannhaupt G., Schwarze C., Vetter I.;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 1487-1835 FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=93377417; PubMed=8356014;
 RA Busseau F., Maillet L., Gallion L., Jaquet M.;
 RL "A 12.8 kb segment, on the right arm of chromosome II from
 RT *Saccharomyces cerevisiae* including part of the DUR1,2 gene, contains
 RT five putative new genes."
 RL Yeast 9:797-806(1993).
 CC
 CC -1- FUNCTION: HYDROLYSIS OF UREA TO AMMONIA AND CO2.
 CC -1- CATALYTIC ACTIVITY: ATP + urea + CO(2) = ADP + phosphate + urea-1-
 CC carboxylate.
 CC -1- CATALYTIC ACTIVITY: Urea-1-carboxylate + H(2)O = 2 CO(2) + 2
 CC NH(3).
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: ALANTOIN AND ARGININE METABOLISM.
 CC -1- SUBUNIT: Monomer.
 CC -1- INDUCTION: BY ALLOPHANATE OR ITS NON-METABOLIZED ANALOG
 CC OXALURATE. REPRESSED IN THE PRESENCE OF READILY USED NITROGEN
 CC SOURCES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M64926; AAC41643.1;
 DR EMBL; Z36077; CAA85172.1;
 DR EMBL; Z21487; CAA79695.1;
 DR PIR; S46082; S46082.
 DR HSSP; P24182; IENC.
 DR SGD; S0000412; DUR1,2.
 DR GO; GO:0004039; F:allophanate hydrolase activity; IMP.
 DR GO; GO:0004847; F:urea carboxylase activity; IMP.
 DR GO; GO:0019627; P:urea metabolism; IMP.
 DR InterPro; IPR000120; Amidase.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR005482; Biotin_carb_C.
 DR InterPro; IPR000089; Biotin_lipoyL.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR005481; CPase_L_N.
 DR InterPro; IPR003778; DUF183.
 DR InterPro; IPR003833; DUF213.
 DR Pfam; PF02682; AHS1; 1.
 DR Pfam; PF02626; AHS2; 1.
 DR Pfam; PF01425; Amidase; 1.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF00364; Biotin_lipoyL; 1.
 DR Pfam; PF00289; CPase_L_D2; 1.
 DR Pfam; PF02786; CPase_L_D2; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPASE_1; 1.
 DR PROSITE; PS00867; CPASE_2; 1.
 DR Ligase; Hydrolase; Multifunctional enzyme; Arginine metabolism;

KW Biotin: ATP-binding. ATP (POTENTIAL).
 FT NP BIND 122 129
 FT BINDING 1798 1798
 FT CONFLICT 96 96
 FT CONFLICT 256 258
 FT CONFLICT 459 459
 FT CONFLICT 830 830
 FT CONFLICT 1395 1395
 FT SEQUENCE 1835 AA; 201831 MW; F52BDDDFE42CD65 CRC64;
 SQ
 Query Match 40.9%; Score 45; DB 1; Length 1835;
 Best Local Similarity 42.1%; Pred. No. 70;
 Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 PSYSFIRADSEVODLRN 20
 DB 102 PSFAYEPSKSKYVELLRN 120

RESULT 10
 BACC_BACLI STANDARD; PRT; 6359 AA.
 ID BACC_BACLI
 AC 068008;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bactracin synthetase 3 (BA3) [includes: ATP-dependent isoleucine
 adenylation (IleA) (isoleucine acetylase); ATP-dependent histidine
 adenylation (D-PheA) (D-phenylalanine acetylase); ATP-dependent
 adenylation (HisA) (histidine acetylase); ATP-dependent aspartate
 adenylation (D-AspA) (D-aspartate acetylase); Aspartate racemase
 adenylation (AspA) (aspartate acetylase); Aspartate racemase
 (EC 5.1.1.13); Phenylalanine racemase [ATP hydrolyzing]
 DE (EC 5.1.1.13); Phenylalanine racemase [ATP hydrolyzing]
 DE BACC.
 GN Bacillus licheniformis.
 OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OK NCBI_TaxId=1402;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10716;
 RX MEDLINE=98089193; PubMed=9427658;
 RA Konz D., Kien A., Schoergerdoffer K., Marahiel M.A.;
 RT "The bacitracin biosynthesis operon of Bacillus licheniformis ATCC
 10716: molecular characterization of three multi-modular peptide
 synthetases.";
 RL Chem. Biol. 4:927-937(1997).
 CC -1- FUNCTION: INDUCES PEPTIDE SYNTHESIS, ACTIVATES AND INCORPORATES
 FIVE AMINO ACIDS, FORMS A THIAZOLINE RING BETWEEN THE FIRST TWO
 AMINO ACIDS AND INCORPORATES A D-GLUTAMINE IN THE FOURTH POSITION.
 CC -1- CATALYTIC ACTIVITY: L-aspartate = D-aspartate.
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine = AMP + diphosphate + D-
 phenylalanine.
 CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 (POTENTIAL).
 CC -1- PATHWAY: Cyclic peptide antibiotic bacitracin biosynthesis.
 CC -1- SUBUNIT: LARGE MULTISUBUNIT COMPLEX OF BAI, BAI2 AND BAI3.
 CC -1- DOMAIN: CONSISTS OF FIVE MODULES WITH TWO EPIMERIZATION DOMAINS IN
 THE SECOND AND FOURTH MODULES, AND A PUTATIVE C-TERMINAL
 THIOESTERASE DOMAIN. EACH MODULE INCORPORATES ONE AMINO ACID INTO
 THE PEPTIDE PRODUCT AND CAN BE FURTHER SUBDIVIDED INTO DOMAINS
 RESPONSIBLE FOR SUBSTRATE ADENYLATION, THIOALATION, CONDENSATION
 (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND
 N-METHYLATION (OPTIONAL).
 CC -1- MISCELLANEOUS: BACTIRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC
 DODECAPETIDES, THAT DIFFER BY ONE OR TWO AMINO ACIDS. THE MOST
 ABUNDANT IS BACTIRACIN A, A BRANCHED CYCLIC DODECAPETIDE. IT
 CONTAINS AN N-TERMINAL LINEAR PENTAPEPTIDE MOIETY (Ile-Cys-Leu-D-
 Glu-Ile) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION
 PRODUCT AND A C-TERMINAL HEPTAPEPTIDE RING (Iys-D-Orn-Ile-D-Phe-
 His-D-Asp-Asn), IN WHICH THE FREE ALPHA-CARBOXY GROUP OF THE C-
 TERMINAL ASN IS BOUND TO THE EPSILON-AMINO GROUP OF LYSINE. IT
 CONTAINS FOUR AMINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7,

CC PHE-9, AND ASP-11).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 acyl carrier domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF007865; AAC06348.1; -
 DR PIR: T31679; T31679.
 DR HSSP: P14687; IAMU.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatn.
 DR InterPro: IPR006163; Pp_bind.
 DR InterPro: IPR006162; Ppantne_attach.
 DR InterPro: IPR001031; Thioesterase.
 DR InterPro: IPR00501; AMP-binding; 5.
 DR Pfam: PF00668; Condensation; 7.
 DR Pfam: PF00550; Pp-binding; 5.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 4.
 DR PROSITE: PS00455; AMP BINDING; 5.
 DR PROSITE: PS00075; ACP DOMAIN; 5.
 KW Ligase; isomerase; Hydrolase; Multifunctional enzyme; Repeat.
 KW Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT REPEAT 461 1034
 FT REPEAT 1517 2064
 FT REPEAT 2999 3570
 FT REPEAT 4047 4612
 FT REPEAT 5549 6129
 FT DOMAIN 966 1033
 FT DOMAIN 1998 2064
 FT DOMAIN 3502 3569
 FT DOMAIN 4544 4610
 FT DOMAIN 6052 6119
 FT BINDING 996 996
 FT BINDING 2028 2028
 FT BINDING 3532 3532
 FT BINDING 4574 4574
 FT BINDING 6082 6082
 SQ SEQUENCE 6359 AA; 722923 MW; 82A273C546253074 CRC64;
 Query Match 40.9%; Score 45; DB 1; Length 6359;
 Best Local Similarity 45.5%; Pred. No. 2.7e+02;
 Matches 10; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 1 VPSYFIRADSEVODLRN 22
 DB 3186 VPSFSF---DSVDFITTL 3203

RESULT 11
 YIAJ_ECOLI STANDARD; PRT; 282 AA.
 ID YIAJ_ECOLI
 AC P37671;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical transcriptional regulator yiaJ.
 DE YIAJ OR B3574.
 GN Escherichia coli.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OK NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;

RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS. STRONG, TO H.INFLUENZAE H11032.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collabora-
 CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U00039; AAB18551.1; -
 CC EMBL: AE000435; AAC76598.1; -
 CC PIR: S47795; S47795.
 CC Ecogen: EG12278; YIAJ.
 CC InterPro: IPR005471; HTH_ICLR.
 CC InterPro: IPR005473; HTH_ICLRlike.
 CC Pfam: PF01614; ICLR.1.
 CC SMART: SM00346; HTH_ICLR.1.
 CC PROSITE: PS01051; HTH_ICLR_FAMILY.1.
 CC DR Hypothetical protein; Transcription regulation; DNA-binding;
 CC KW Complete proteome.
 CC FT DNA BIND 45 64 H-T-H MOTIF (POTENTIAL).
 CC SEQUENCE 282 AA; 31066 MW; 228887B672B958F CRC64;
 SQ
 Query Match 40.0%; Score 44; DB 1; Length 282;
 Best Local Similarity 43.5%; Pred. No. 13;
 Matches 10; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
 QY 2 PSY-SFIRHSEVQDLIRNTI 22
 DB 169 PDVKSVMESHQHEIPLIRNTI 191
 RESULT 12
 M12D_BACSU STANDARD; PRT; 344 AA.
 AC P26935;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Myo-inositol 2-dehydrogenase (EC 1.1.1.18).
 GN IDH OR IDLG OR E83G.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCB1_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / 60015;
 RA MEDLINE=92104493; PubMed=1761221.
 RA Fujita Y., Shindo K., Miwa Y., Yoshida K.,
 RT "Bacillus subtilis inositol dehydrogenase-encoding gene (idh):
 RT sequence and expression in Escherichia coli.";
 RL Gene 108:121-125(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSC1A1;
 RA MEDLINE=95039891; PubMed=7952181;
 RA Yoshida K., Sano H., Miwa Y., Ogasawara N., Fujita Y.,
 RT "Cloning and nucleotide sequencing of a 15 kb region of the Bacillus
 RT subtilis genome containing the idh operon.";
 RL Microbiology 140:2289-2298(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=98044033; PubMed=9384377;
 RA Kunet F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Bourstier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capano V., Carter N.M.,
 RA Choi S.K., Codani J.M., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Darioz F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Goltz E.J., Grandi G.,
 RA Gilm S.Y., Glaeser P., Gottreau A., Halesch U., Harwood C.R., Henaut A.,
 RA Giuseppe G., Guy B.J., Haga K., Halech U., Harwood C.R., Henaut A.,
 RA Hilbert H., Holappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manel C., Medigue C.,
 RA Medina N., Meliado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porvolik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro F., Shin B.G., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassartotti A.,
 RA Viart A., Wandut R., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=79239346; PubMed=112095;
 RA Ramsey R., Fujita Y., Freese E.;
 RT "Purification and properties of Bacillus subtilis inositol
 RT dehydrogenase.";
 RL J. Biol. Chem. 254:7684-7690(1979).
 CC -1- CATALYTIC ACTIVITY: Myo-inositol + NAD(+) = 2,4,6/3,5-
 CC pentahydroxycyclohexanone + NADH.
 CC -1- PATHWAY: Myo-inositol catabolism; first step.
 CC -1- SUBUNIT: Monomer.
 CC -1- INDUCTION: INOSITOL, SUBJECTED TO CATABOLITE REPRESSION.
 CC -1- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collabora-
 CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M76431; AAA22543.1; -
 CC EMBL: D14399; BAA03296.1; -
 CC EMBL: Z99124; CAB16006.1; -
 CC PIR: JH0511; JH0511; idh.
 CC Subtilist; BG10669; idh.
 CC InterPro: IPR000683; GFO_IDH_MOCA.
 CC InterPro: IPR004104; GFO_IDH_MOCA_C.
 CC Pfam: PF01408; GFO_IDH_MOCA.1.
 CC Pfam: PF02894; GFO_IDH_MOCA_C.1.
 CC DR Oxidoreductase, NAD, Complete proteome.
 CC KW SEQUENCE 344 AA; 38351 MW; 2FCB908D4E2C332F CRC64;
 SQ
 Query Match 40.0%; Score 44; DB 1; Length 344;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 6 FIRHSEVQDLIRNTI 21
 DB 276 FVAAYDVEIOPFIDSI 291
 RESULT 13
 YN04_YEAST

```

ID YN14 YEAST STANDARD; PRT; 587 AA.
AC P53933.1
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 66.1 kDa protein in RPS7B-YPT53 intergenic region.
GN YN1094W OR N2219.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9288C / FY1679;
RA MEDLINE=96367601; PubMed=8771715;
RA Garcia-Cantalejo J.M., Boskovic U., Jimenez A.;
RT "Sequence analysis of a 14.2 kb fragment of Saccharomyces cerevisiae
RT chromosome XIV that includes the YPT53, ERN1 and gsr m2 genes and
RT four new open reading frames.";
RL Yeast 12:599-608(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X85811; CAAS9823.1; ALT_INT.
DR EMBL: Z71370; CA95970.1; -.
DR PIR: S63033; S63033.
DR SGD: S0005038; YN1094W.
DR GO: GO:0015629; C:actin cytoskeleton. IPI.
DR GO: GO:0030036; P:actin cytoskeleton organization and biogenesis; IPI.
DR Hypothetical protein.
KW SEQUENCE 587 AA; 66134 MW; B85C525548BA34BC CRC64;
SQ
Query Match 40.0%; Score 44; DB 1; Length 587;
Best Local Similarity 45.0%; Pred. No. 29;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 3 SYSFIRADSEYVDLIRNII 22
DB 430 SMSDVPSHDEYMEVNNII 449

```

```

CC -1- FUNCTION: PHOSPHORYLATES PTDINS. PTDINS4P AND PTDINS(4,5)P2 WITH A
CC PREFERENCE FOR PTDINS(4,5)P2.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-D-myo-inositol 4,5-
CC bisphosphate = ADP + 1-phosphatidy1-D-myo-inositol 3,4,5-
CC trisphosphate.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M93252; AAA30698.1; -.
DR PIR: A43322; A43322.
DR InterPro: IPR000008; C2.
DR InterPro: IPR000403; PI3_P14_Kinase.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR003113; PI3K_P85B.
DR InterPro: IPR000341; PI3K_ras_Bind.
DR InterPro: IPR01263; PI3Ka.
DR Pfam: PF00454; PI3_P14_Kinase; 1.
DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF02192; PI3K_P85B; 1.
DR Pfam: PF00794; PI3K_P85B; 1.
DR Pfam: PF00613; PI3Ka; 1.
DR SMART: SM00239; C2; 1.
DR SMART: SM00142; PI3K_C2; 1.
DR SMART: SM00143; PI3K_P85B; 1.
DR SMART: SM00144; PI3K_P85B; 1.
DR SMART: SM00145; PI3Ka; 1.
DR SMART: SM00146; PI3K_P85B; 1.
DR PROSITE: PS50004; C2 DOMAIN 2; FALSE_NEG.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS50290; PI3_4_KINASE_3; 1.
DR TRANSFERASE; Kinase; Multigene family.
KW DOMAIN 319 428 PI3K/PI4K
FT DOMAIN 797 1068 PI3K/PI4K
SQ SEQUENCE 1068 AA; 124327 MW; C753DC2F39FDDF0 CRC64;
Query Match 40.0%; Score 44; DB 1; Length 1068;
Best Local Similarity 45.5%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 1 VPSYFIRADSEYVDLIRNII 22
DB 123 MVECFEPMVKDPEVDFRRNII 144

```

```

RESULT 14
P11A_BOVIN STANDARD; PRT; 1068 AA.
ID P11A_BOVIN
AC P32871.1
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit.
DE alpha isoform (EC 2.7.1.153) (PI3-kinase p110 subunit alpha) (PcdIns-
DE 3-kinase p110) (PI3K).
GN PIK3CA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92354059; PubMed=1322797;
RA Hiles I.D., Otsu M., Volinia S., Fry M.J., Gout I., Dhand R.,
RA Panayotov G., Ruiz-Larrea F., Thompson A., Totty N.F., Hsuan J.J.,
RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
RA "Phosphatidylinositol 3-kinase: structure and expression of the 110
RT kd catalytic subunit.";
RL Cell 70:419-429(1992).

```

```

RESULT 15
P11A_HUMAN STANDARD; PRT; 1068 AA.
ID P11A_HUMAN
AC P42336.1
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit.
DE alpha isoform (EC 2.7.1.153) (PI3-kinase p110 subunit alpha) (PcdIns-
DE 3-kinase p110) (PI3K).
GN PIK3CA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95229146; PubMed=7713498;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:31:40 ; Search time 33.1564 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049A-15
Perfect score: 110
Sequence: 1 VPSYFIRAHDSVEVDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHASE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	98	89.1	1590	2	059983	059983 streptococc
2	98	89.1	1590	2	055263	055263 streptococc
3	83	75.5	1290	2	048756	048756 leuconostoc
4	83	75.5	1477	2	091466	091466 leuconostoc
5	83	75.5	1508	2	09EZH5	09EZH5 leuconostoc
6	83	75.5	1508	2	052224	052224 leuconostoc
7	82	74.5	1527	2	092AR4	092AR4 leuconostoc
8	82	74.5	1527	2	08KRE1	08KRE1 leuconostoc
9	82	74.5	1554	2	08KJ5	08KJ5 streptococc
10	78	70.9	2835	2	08G9Q2	08G9Q2 leuconostoc
11	78	70.9	2835	2	09WJ5	09WJ5 streptococc
12	76	69.1	1338	2	09WJ4	09WJ4 streptococc
13	76	69.1	1338	2	08S42	08S42 streptococc
14	71	64.5	1449	2	055264	055264 streptococc
15	71	64.5	1449	2	09RE05	09RE05 leuconostoc
16	71	64.5	2057	2	09RE05	09RE05 leuconostoc

17	70	63.6	1577	2	055265	055265 streptococc
18	68	61.8	1518	2	000600	000600 streptococc
19	68	61.8	1575	2	091CH3	091CH3 streptococc
20	68	61.8	1577	2	054178	054178 streptococc
21	68	61.8	1599	2	000599	000599 streptococc
22	50	45.5	51	16	09RNP7	09RNP7 vibrio chol
23	50	45.5	93	2	0921X9	0921X9 methanosarc
24	48	43.6	244	17	0800D3	0800D3 giardia lam
25	47	42.7	480	5	095WU6	095WU6 bacillus sp
26	47	42.7	639	2	09AQR3	09AQR3 cryptococcu
27	47	42.7	1554	3	08U026	08U026 cryptococcu
28	47	42.7	1567	3	08J0M2	08J0M2 cryptococcu
29	46.5	42.3	414	10	065491	065491 arabidopsis
30	46	41.8	66	2	09XSK6	09XSK6 bacillus mo
31	46	41.8	200	16	08XJ47	08XJ47 clostridium
32	46	41.8	326	12	08OV16	08OV16 avian reovi
33	46	41.8	597	16	09RPS8	09RPS8 ureaplasma
34	46	41.8	631	16	09RPL8	09RPL8 mycoplasma
35	46	41.8	1032	10	09F1J9	09F1J9 arabidopsis
36	46	41.8	1035	10	08GS71	08GS71 rhizobium
37	45.5	41.4	248	16	0987H1	0987H1 schizosacch
38	45.5	41.4	767	3	074759	074759 clostridium
39	45.5	41.4	956	16	09PJP0	09PJP0 clostridium
40	45	40.9	111	16	08XLB3	08XLB3 bacterioph
41	45	40.9	136	9	064072	064072 bacterioph
42	45	40.9	136	16	031949	031949 bacillus su
43	45	40.9	326	12	08OV08	08OV08 avian reovi
44	45	40.9	852	13	09PTES	09PTES xenopus lae
45	44.5	40.5	462	2	08GIY5	08GIY5 mycoplasma

ALIGNMENTS

RESULT 1	ID	Query Match	Score	DB	Length	Matches	Indels	Gaps
059983	PRELIMINARY;	PRT;	1590	AA.				
AC	059983;							
DT	01-NOV-1996 (TREMELREL. 01, Created)							
DT	01-NOV-1996 (TREMELREL. 01, Last sequence update)							
DT	01-OCT-2002 (TREMELREL. 22, Last annotation update)							
DE	Glucosyltransferase-I precursor (EC 2.4.1.5).							
GN	GTFI.							
OS	Streptococcus sobrinus.							
OC	Bacteri; Firmicutes; Lactobacillales; Streptococcaceae;							
OX	Streptococcus.							
OX	NCBI_TaxID=1310;							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=OM2176;							
RX	MEDLINE=94146405; PubMed=8312602;							
RA	Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;							
RT	"DNA sequence of the glucosyltransferase gene of serotype d							
RT	Streptococcus sobrinus.";							
RL	DNA Seq. 4:19-27(1993).							
DR	EMBL: D13858; BAA02976.1; -;							
DR	InterPro: IPR002479; CM binding.							
DR	InterPro: IPR003318; Glyco_hydro_70.							
DR	Pfam: PF01473; CM_binding_1; 16.							
DR	Pfam: PF02324; Glyco_hydro_70; 1.							
KW	Glycosyltransferase; Signal; Transferase.							
FT	SIGNAL 1							
FT	CHARIN 39 1590							
FT	GLUCOSYLTRANSFERASE-I.							
FT	SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;							
QY	1 VPSYFIRAHDSVEVDLIRNII 22							
DB	548 VPSYFIRAHDSVEVDLIRNII 569							

RESULT 2

Q5263 PRELIMINARY; PRT; 1590 AA.
 ID Q5263
 AC Q5263
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 SC Streptococcus.
 OX NCBI_TaxID=1310;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase
 produced from Streptococcus sobrinus ATCC 33478."
 RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 15.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3FC6B4FD43 CRC64;

Query Match 89.1%; Score 98; DB 2; Length 1590;
 Best Local Similarity 86.4%; Pred. No. 8.1e-07;
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYFTRAHDSVQDLIRNII 22
 |||||
 DB 548 VPSYFTRAHDSVQDLIRNII 569

RESULT 3

Q48756 PRELIMINARY; PRT; 1290 AA.
 ID Q48756
 AC Q48756
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B1299;
 RA MEDLINE=9713666; PubMed=8982063;
 RA Monchois V., Willemot R.M., Renaud-Simeon M., Croux C., Monsan P.;
 RT "Cloning and sequencing of a gene coding for a novel dextranucrase
 from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-
 6) and alpha (1-3) linkages."
 RT Gene 182:23-32(1996).
 RL EMBL; U38181; AAB40875.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 11.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1290 AA; 145590 MW; 3555C2E96B749FMA CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1290;
 Best Local Similarity 76.2%; Pred. No. 0.00018;
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PYSYFTRAHDSVQDLIRNII 22
 :|||
 DB 388 PYSYFTRAHDSVQDLIRNII 408

RESULT 4

Q9L466 PRELIMINARY; PRT; 1477 AA.
 ID Q9L466
 AC Q9L466
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRG.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon M., Pizzut S., Sarcabal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
 glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A0250172; CAB76565.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Glycoyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFCB831 CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1477;
 Best Local Similarity 68.2%; Pred. No. 0.0002;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFTRAHDSVQDLIRNII 22
 :|||
 DB 603 IPYSFVRAHDSVQDLIRNII 624

RESULT 5

Q9EZH5 PRELIMINARY; PRT; 1508 AA.
 ID Q9EZH5
 AC Q9EZH5
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase DsrB742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-742CB;
 RA Kim H.-S., Kim D., Ryu H.-J., Rodye J.F.;
 RA "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF294469; AAG38021.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87A4F3A CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.00021;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFTRAHDSVQDLIRNII 22
 :|||
 DB 634 IPYSFVRAHDSVQDLIRNII 655

RESULT 6

Q52224

ID 052224 PRELIMINARY; PRT; 1508 AA.
 AC 052224;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Remaud-Simeon M., Monsan P., Willemot R.M.;
 RT "Cloning and sequencing of a gene coding for an extracellular
 dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
 RT synthesizing only a (1-6) glucan."
 RL EMBL; AF030129; AAB95453.1; -
 DR FEMS Microbiol. Lett. 0:0-0(1998).
 DR InterPro: IPR002479; Glyco_hydro_70.
 DR InterPro: IPR003318; Glyco_hydro_14.
 DR Pfam; PF01473; CM_binding_1; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70DD1F0 CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.00021;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 634 IPNYSFVRADSEVQTVIAQII 655

RESULT 7
 ID 092AR4 PRELIMINARY; PRT; 1527 AA.
 AC 092AR4;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DEX.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512-F;
 RA Bhanagar R., Singh D.K.S.;
 RT "Cloning and Molecular Characterization of Dextranucrase Gene from
 RT Leuconostoc mesenteroides NRRL B-512F."
 RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U81374; AAD10952.1; -
 DR InterPro: IPR002479; CM_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1527 AA; 169709 MW; IDFAFA237C743398 CRC64;

Query Match 74.5%; Score 82; DB 2; Length 1527;
 Best Local Similarity 63.6%; Pred. No. 0.00031;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 652 IPNYSFVRADSEVQTVIAQIV 673

RESULT 8
 Q8KREI PRELIMINARY; PRT; 1527 AA.
 ID Q8KREI

AC Q8KREI;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Dextranucrase DsrD (EC 2.4.1.5).
 GN DSRD.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RA Neubauer H., Bauche A., Mollet B.;
 RT "Isolation and characterization of the dextranucrase DsrD of
 RT Leuconostoc mesenteroides Lcc4."
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY017384; AAG61158.1; -
 DR InterPro: IPR002479; CM_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferase; Glycosyltransferase.
 KM SEQUENCE 1527 AA; 169835 MW; F9DDDE220BD89668 CRC64;

Query Match 74.5%; Score 82; DB 2; Length 1527;
 Best Local Similarity 63.6%; Pred. No. 0.00031;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 652 IPNYSFVRADSEVQTVIAQIV 673

RESULT 9
 ID 08KZL5 PRELIMINARY; PRT; 1554 AA.
 AC 08KZL5;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 NCBI_TaxID=1310;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; Pubmed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan."
 RL Biochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL; AB089438; BAC07265.1; -
 DR InterPro: IPR002479; CM_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferase.
 KM SEQUENCE 1554 AA; 171676 MW; 6981BCC1DAE24A73 CRC64;

Query Match 74.5%; Score 82; DB 2; Length 1554;
 Best Local Similarity 63.6%; Pred. No. 0.00031;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 557 IPNYSFVRADSEVQTVIAKIV 578

RESULT 10
 Q9LCJ7

ID 09LCU7 PRELIMINARY; PRT; 1016 AA.
 AC 09LCU7;
 GN 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DSRT.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RX NCBI_TaxID=1245;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512F;
 RX MEDLINE=20169623; PubMed=10705445; Kobayashi M.;
 RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
 RT "gene encoding a dextranucrase-like protein in Leuconostoc
 RT mesenteroides NRRL B-512F.";
 RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
 DR EMBL; AB020020; BAA90527.1; -.
 DR HSSP; P06278; 1VTS.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1016 AA; 110344 MW; 8896FDEB13CCB47 CRC64;
 QY Query Match 72.7%; Score 80; DB 2; Length 1016;
 Best Local Similarity 71.4%; Pred. No. 0.00041;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PSYSEFRAHDESVQDLIRNII 22
 DB 625 PNYAFIRAHDESVQTVIAQII 645

RESULT 11
 Q8G9Q2 PRELIMINARY; PRT; 2835 AA.
 ID 08G9Q2;
 AC 08G9Q2;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DE Dextranucrase (EC 2.4.1.5) (Fragment).
 GN DSRE.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RX NCBI_TaxID=1245;
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22231661; PubMed=12270834;
 RA Bozonnet S., Dols-Lafargue M., Fabre E., Pizut S., Renaud-Simeon M.,
 RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
 RT synthesizing dextranucrase with two catalytic domains.";
 RL J. Bacteriol. 184:5753-5761(2002).
 DR EMBL; AJ430204; CAD22883.1; -.
 DR Transferase; Glycosyltransferase.
 GN NON TER
 FT SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;
 QY Query Match 70.9%; Score 78; DB 2; Length 2835;
 Best Local Similarity 71.4%; Pred. No. 0.00277;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PSYSEFRAHDESVQDLIRNII 22
 DB 629 PNYAFIRAHDESVQTVIAQII 649

RESULT 12
 Q9WXJ5 PRELIMINARY; PRT; 1512 AA.
 ID 09WXJ5;
 AC 09WXJ5;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF.
 OS Streptococcus criceti.
 OG Plasmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NCBI_TaxID=1333;
 RX NCBI_TaxID=1333;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S. cricetus glucosyltransferase (gifs and gtfS) genes.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB026123; BAA77237.1; -.
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Plasmid.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;
 QY Query Match 70.0%; Score 77; DB 2; Length 1512;
 Best Local Similarity 71.4%; Pred. No. 0.002;
 Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 PSYSEFRAHDESVQDLIRNII 22
 DB 560 PNYAFIRAHDESVQTVIAQII 580

RESULT 13
 Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
 ID 09WXJ4;
 AC 09WXJ4;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF-S.
 OS Streptococcus criceti.
 OG Plasmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NCBI_TaxID=1333;
 RX NCBI_TaxID=1333;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S. cricetus glucosyltransferase (gifs and gtfS) genes.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB026123; BAA77237.1; -.
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 10.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Plasmid.
 SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;
 QY Query Match 69.1%; Score 76; DB 2; Length 1338;
 Best Local Similarity 72.7%; Pred. No. 0.0025;
 Matches 16; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VPSYSEFRAHDESVQDLIRNII 22
 DB 509 PNYAFIRAHDESVQTVIAQII 530

RESULT 14
 O68542 PRELIMINARY; PRT; 1449 AA.
 ID O68542;
 AC O68542;
 DT 01-AUG-1998 (Tremblrel. 07, Created)


```

DT 01-AUG-1998 (TRENBLREL. 07, last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 gTfN."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase.
FT NON_TER 1449 1449
SQ SEQUENCE 1449 AA; 159895 MW; 0700FED748471BFB CRC64;

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.018;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVODLRNII 22
Db 609 NYAFVRAHDSVOSIIIGI 628

RESULT 15
Q55264 PRELIMINARY; PRT; 1449 AA.
ID Q55264
AC Q55264;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, last annotation update)
DE Glucosyltransferase precursor.
GN GTFP.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases."
RL Infect. Immun. 63:609-621 (1995).
DR EMBL; U35495; AAC41412.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Signal; Transferase.
FT SIGNAL 1 35
FT CHAIN 36 1449 POTENTIAL.
SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.018;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVODLRNII 22
Db 609 NYAFVRAHDSVOSIIIGI 628

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 13, 2003, 09:11:40 ; Search time 14.2844 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-15
Perfect score: 110
Sequence: 1 VPSYSFIRAHDSVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 segs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB pep:*
5: /cgn2_6/prodata/2/iaa/PCUTS COMB pep:*
6: /cgn2_6/prodata/2/iaa/backfile1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	1375	3	US-09-210-361-4
2	110	100.0	1375	4	US-09-740-274-4
3	99	90.0	1475	3	US-09-007-999-2
4	99	90.0	1475	3	US-09-210-361-2
5	99	90.0	1475	3	US-09-740-274-2
6	82	74.5	523	4	US-09-604-957-5
7	76	69.1	535	4	US-09-604-957-7
8	76	69.1	1278	4	US-09-604-957-3
9	71	64.5	2057	4	US-09-499-203-2
10	70	63.6	1577	2	US-08-793-824-2
11	68	61.8	545	4	US-09-604-957-4
12	68	61.8	1430	3	US-09-008-172-2
13	68	61.8	1430	3	US-09-210-361-6
14	68	61.8	1430	3	US-09-740-274-6
15	66	60.0	584	4	US-09-604-957-6
16	50	45.5	565	4	US-09-107-532A-4217
17	47	42.7	639	4	US-09-509-814A-4
18	46	41.8	196	2	US-08-684-024-2
19	46	41.8	196	2	US-08-684-024-2
20	46	41.8	196	3	US-09-145-868-2
21	46	41.8	196	3	US-09-145-868-9
22	44	40.0	282	4	US-09-172-952-20
23	44	40.0	1068	3	US-08-390-874C-11
24	44	40.0	1068	3	US-09-265-772-11
25	44	40.0	1069	2	US-08-162-081B-37
26	44	40.0	1069	2	US-08-780-872-37
27	44	40.0	1069	3	US-09-085-957-37

28	44	40.0	1090	2	US-08-162-081B-36	Sequence 36, Appl
29	44	40.0	1080	2	US-08-780-872-36	Sequence 36, Appl
30	44	40.0	1080	3	US-09-085-957-36	Sequence 36, Appl
31	43	39.1	75	4	US-09-107-532A-7020	Sequence 2700, Ap
32	43	39.1	347	4	US-09-461-325-277	Sequence 4975, Ap
33	42	38.2	211	4	US-09-328-352-4995	Sequence 277, Ap
34	42	38.2	2627	2	US-08-751-189-3	Sequence 3, Appl
35	42	38.2	2627	2	US-09-060-836-3	Sequence 3, Appl
36	42	38.2	2627	3	US-09-184-445-3	Sequence 31249, A
37	41.5	37.7	348	4	US-09-252-991A-31249	Sequence 4, Appl
38	41.5	37.7	713	2	US-08-849-212-4	Sequence 38, Appl
39	41	37.3	616	4	US-09-268-347-38	Sequence 46, Appl
40	41	37.3	659	4	US-09-268-347-46	Sequence 45, Appl
41	41	37.3	660	4	US-09-268-347-45	Sequence 37, Appl
42	40	36.4	309	4	US-09-345-473E-37	Sequence 10, Appl
43	40	36.4	472	2	US-08-272-255-10	Sequence 4, Appl
44	40	36.4	472	3	US-08-964-268-4	Sequence 4, Appl
45	40	36.4	472	4	US-09-105-254-4	Sequence 4, Appl

ALIGNMENTS

```

RESULT 1
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match          100.0%; Score 110; DB 3; Length 1375;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 VPSYSFIRAHDSVQDLIRNII 22
 Db 578 VPSYSFIRAHDSVQDLIRNII 599

```

RESULT 2
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19

```

```

1 PRIOR APPLICATION NUMBER: 09/210,361
2 PRIOR FILING DATE: 1998-12-11
3 PRIOR APPLICATION NUMBER: 09/007,999
4 PRIOR FILING DATE: 1998-01-16
5 PRIOR APPLICATION NUMBER: 08/478,704
6 PRIOR FILING DATE: 1995-06-07
7 PRIOR APPLICATION NUMBER: 09/009,620
8 PRIOR FILING DATE: 1998-01-20
9 PRIOR APPLICATION NUMBER: 08/485,243
10 PRIOR FILING DATE: 1995-06-07
11 PRIOR APPLICATION NUMBER: 09/008,172
12 PRIOR FILING DATE: 1998-01-16
13 PRIOR APPLICATION NUMBER: 08/482,711
14 PRIOR FILING DATE: 1995-06-07
15 SOFTWARE: FastSeq for Windows Version 3.0
16 SEQ ID NO: 4
17 LENGTH: 1375
18 TYPE: PR1
19 ORGANISM: streptococcus mutans
20 US-09-740-274-4

```

```

Query Match      100.0%; Score 110; DB 4; length 1375;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 VPSYSFIRAHDSVQDLIRNII 22
        |||||
Db      578 VPSYSFIRAHDSVQDLIRNII 599
```

```

RESULT 3
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: Latexes in Paper Manufacture
; FILE REFERENCE: 0355D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Fastrsed for Windows Version 3.0
; SEQ ID NO. 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

```

Query Match	90.0%;	Score 99;	DB 3;	Length 1475;
Best Local Similarity	90.9%;	Pred. No. 2.9e-08;		
Matches 20;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;

```

QY      1 VPSYSFIRAHDSVQDLIRNII 22
        |||||
Db      552 VPSYSFIRAHDSVQDLIADII 573

```

RESULT 4
US-09-210-361-2
Sequence 2, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
TITLE OF INVENTION: Latexes in Paper Manufacture
FILE REFERENCE: 035/0CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999

```

? EARLIER FILING DATE: 1998-01-16
? EARLIER APPLICATION NUMBER: 08/478,704
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 09/009,620
? EARLIER FILING DATE: 1998-01-20
? EARLIER APPLICATION NUMBER: 08/485,243
? EARLIER FILING DATE: 1995-06-07
? EARLIER APPLICATION NUMBER: 09/008,172
? EARLIER FILING DATE: 1998-01-16
? EARLIER APPLICATION NUMBER: 08/482,711
? EARLIER FILING DATE: 1995-06-07
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 2
? LENGTH: 1475
? TYPE: prt
? ORGANISM: Streptococcus mutans
? US-09-210-361-2

```

Query Match	90.0%;	Score 99;	DB 3;	Length 1475;
Best Local Similarity	90.9%;	Pred. No. 2.9e-08;		
Matches 20; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      1 VPSYSFIRAHDSVQDLIRNI 22  
         ||||| : ||  
Db      552 VPSYSFIRAHDSVQDLADI 573
```

```

RESULT 5
US-09-740-274-2
/ Sequence 2, Application US/09740274
/ Patent No. 6465203
/ GENERAL INFORMATION:
/ APPLICANT: Nichols, Scott E.
/ TITLE OF INVENTION: Glucan-containing Compositions and Paper
/ FILE REFERENCE: 0357CRD
/ CURRENT APPLICATION NUMBER: US/09/740,274
/ CURRENT FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 09/210,361
/ PRIOR FILING DATE: 1998-12-11
/ PRIOR APPLICATION NUMBER: 09/007,999
/ PRIOR FILING DATE: 1998-01-16
/ PRIOR APPLICATION NUMBER: 08/478,704
/ PRIOR FILING DATE: 1995-06-07
/ PRIOR APPLICATION NUMBER: 09/009,620
/ PRIOR FILING DATE: 1998-01-20
/ PRIOR APPLICATION NUMBER: 08/485,243
/ PRIOR FILING DATE: 1995-06-07
/ PRIOR APPLICATION NUMBER: 09/008,172
/ PRIOR FILING DATE: 1998-01-16
/ PRIOR APPLICATION NUMBER: 08/482,711
/ PRIOR FILING DATE: 1995-06-07
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 1475
/ TYPE: PR1
/ ORGANISM: Streptococcus mutans
/ US-09-740-274-2

```

Query Match	90.0%;	Score 99;	DB 4;	Length 1475;
Best Local Similarity	90.9%;	Pred. No. 2.9e-08;		
Matches	20;	Conservative	1;	Mismatches 1; Indels 0; Gaps 0;

```
QY      1 VPSYSFIRAHDSVEVDLIRNII 22  
        |||||  
Db      552 VPSYSFIRAHDSVEVDLIADII 573
```

RESULT 6
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314

```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match          74.5%; Score 82; DB 4; Length 523;
Best Local Similarity 63.6%; Pred. No. 6.8e-06;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDLIRNII 22
DB 146 IPNYSFVRAHDSVQDLIRNII 167

RESULT 7
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

Query Match          69.1%; Score 76; DB 4; Length 535;
Best Local Similarity 59.1%; Pred. No. 7.4e-05;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDLIRNII 22
DB 144 IPNYSFVRAHDSVQDLIRNII 165

RESULT 8
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
```

```

; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

Query Match          69.1%; Score 76; DB 4; Length 1278;
Best Local Similarity 59.1%; Pred. No. 0.00021;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDLIRNII 22
DB 620 IPNYSFVRAHDSVQDLIRNII 641

RESULT 9
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match          64.5%; Score 71; DB 4; Length 2057;
Best Local Similarity 63.2%; Pred. No. 0.0027;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDLIR 19
DB 757 IPNYSFVRAHDVDAQDLIR 775

RESULT 10
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2
```

```
Query Match 63.6%; Score 70; DB 2; Length 1577;
Best Local Similarity 65.0%; Pred. No. 0.0029;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 SYSFIRAHSEVQDILRNII 22
: |||||: |||
Db 661 NYIFIRAHSEVQVIAKII 660
```

```
RESULT 11
US-09-604-957-4
Sequence 4, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDIJNA HENDRIKA
APPLICANT: DIJKHUIZEN, LOBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 545
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-604-957-4
```

```
Query Match 61.8%; Score 68; DB 4; Length 545;
Best Local Similarity 70.0%; Pred. No. 0.0018;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 3 SYSFIRAHSEVQDILRNII 22
: |||||: |||
Db 156 NYIFIRAHSEVQVIAKII 175
```

```
RESULT 12
US-09-008-172-2
Sequence 2, Application US/09008172
Patent No. 6127602
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
TITLE OF INVENTION: Latexes in Paper Manufacture
FILE REFERENCE: 0358D
CURRENT APPLICATION NUMBER: US/09/008,172
PRIOR FILING DATE: 1998-01-16
```

```
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1430
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-008-172-2
```

```
Query Match 61.8%; Score 68; DB 3; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.0057;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 3 SYSFIRAHSEVQDILRNII 22
: |||||: |||
Db 576 NYIFIRAHSEVQVIAKII 595
```

```
RESULT 13
US-09-210-361-6
Sequence 6, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
TITLE OF INVENTION: Latexes in Paper Manufacture
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-210-361-6
```

```
Query Match 61.8%; Score 68; DB 3; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.0057;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 3 SYSFIRAHSEVQDILRNII 22
: |||||: |||
Db 576 NYIFIRAHSEVQVIAKII 595
```

```
RESULT 14
US-09-740-274-6
Sequence 6, Application US/09740274
Patent No. 645203
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
```

; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 61.8%; Score 68; DB 4; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.0057;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
Db 576 NTFIRAHSEVQTVIAKII 595

RESULT 15
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUT, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 60.0%; Score 66; DB 4; Length 584;
Best Local Similarity 61.1%; Pred. No. 0.0043;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFIRAHSEVQDLI 18
Db 167 IPNYSFVRAHDYDAQDPI 184

Search Completed: November 13, 2003, 09:48:01
Job time : 14.2844 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 28.7773 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-15

Perfect score: 110
Sequence: 1 VPSYSTRADSEVQDLIRNII 22

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09A_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	1375	9	US-09-740-274-4 Sequence 4, Appl1
2	99	90.0	1475	9	US-09-740-274-2 Sequence 2, Appl1
3	82	74.5	522	10	US-09-995-749A-11 Sequence 11, Appl1
4	76	69.1	535	10	US-09-995-749A-13 Sequence 13, Appl1
5	76	69.1	1781	10	US-09-995-749A-2 Sequence 2, Appl1
6	71	64.5	584	10	US-09-995-749A-12 Sequence 12, Appl1
7	68	61.8	545	10	US-09-995-749A-10 Sequence 10, Appl1
8	68	61.8	1430	9	US-09-740-274-6 Sequence 6, Appl1
9	44	40.0	282	11	US-09-557-796-20 Sequence 20, Appl1
10	44	40.0	1068	12	US-10-101-235A-8 Sequence 8, Appl1
11	43	39.1	347	15	US-10-012-542-277 Sequence 277, App
12	43	39.1	568	15	US-10-156-761-11791 Sequence 11791, A
13	43	39.1	2541	15	US-10-177-293-470 Sequence 470, App
14	41.5	37.7	1213	9	US-09-815-242-5358 Sequence 5358, Ap
15	41.5	37.7	1217	9	US-09-815-242-12523 Sequence 12523, A

16	40.5	36.8	296	14	US-10-080-960-19 Sequence 19, Appl1
17	40.5	36.8	481	9	US-09-815-242-5584 Sequence 5584, Ap
18	40.5	36.8	487	9	US-09-815-242-12456 Sequence 12456, A
19	40	36.4	43	11	US-09-813-153-163 Sequence 163, Appl
20	40	36.4	309	10	US-09-862-027-37 Sequence 37, Appl
21	40	36.4	472	12	US-10-413-536-4 Sequence 4, Appl1
22	40	36.4	489	10	US-09-738-626-3805 Sequence 3805, Ap
23	40	36.4	572	9	US-09-815-242-5626 Sequence 5626, Ap
24	40	36.4	573	9	US-09-815-242-12247 Sequence 12247, A
25	40	36.4	574	9	US-09-815-242-12291 Sequence 12291, A
26	40	36.4	574	9	US-09-815-242-12045 Sequence 12045, A
27	40	36.4	779	12	US-10-353-929-49 Sequence 49, Appl
28	39.5	35.9	443	15	US-10-286-264-106 Sequence 106, App
29	39	35.5	67	15	US-10-007-280A-214 Sequence 214, App
30	39	35.5	108	9	US-09-764-868-895 Sequence 895, App
31	39	35.5	188	9	US-09-925-287-482 Sequence 482, App
32	39	35.5	228	10	US-09-895-913A-126 Sequence 126, App
33	39	35.5	405	15	US-10-156-761-9148 Sequence 9148, Ap
34	39	35.5	557	11	US-09-934-455-326 Sequence 326, App
35	39	35.5	1383	15	US-10-021-955-82 Sequence 82, Appl
36	39	35.5	1383	15	US-10-021-955-86 Sequence 86, Appl
37	39	35.5	1389	15	US-10-021-955-87 Sequence 87, Appl
38	39	35.5	1389	15	US-10-021-955-79 Sequence 79, Appl
39	39	35.5	1714	15	US-10-098-916A-11 Sequence 11, Appl
40	38.5	35.0	92	15	US-10-106-698-7618 Sequence 7618, Ap
41	38.5	35.0	338	11	US-09-769-787-146 Sequence 146, App
42	38.5	35.0	443	12	US-10-205-219-91 Sequence 91, Appl
43	38.5	35.0	456	10	US-09-925-300-1585 Sequence 1585, Ap
44	38.5	35.0	579	15	US-10-156-761-8655 Sequence 8655, Ap
45	38.5	35.0	887	9	US-09-815-242-13179 Sequence 13179, A

ALIGNMENTS

```
RESULT 1
US-09-740-274-4
Sequence 4, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OR INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 100.0%; Score 110; DB 9; Length 1375;
Best local similarity 100.0%; Pred. No. 3, 6e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPSYSTRADSEVQDLIRNII 22
|||||
```

DB 578 VPSYFIRAHDSVQDLIRNII 599

RESULT 2

US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 90.0%; Score 99; DB 9; Length 1475;
Best Local Similarity 90.9%; Pred. No. 2.5e-07;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSVQDLIRNII 22

DB 552 VPSYFIRAHDSVQDLIRNII 573

RESULT 3

US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 74.5%; Score 82; DB 10; Length 522;
Best Local Similarity 63.6%; Pred. No. 4.7e-05;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSVQDLIRNII 22

DB 146 IPNYSFVRAHDNSQDLIAQIV 167

RESULT 4

US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 69.1%; Score 76; DB 10; Length 535;
Best Local Similarity 59.1%; Pred. No. 0.00046;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSVQDLIRNII 22

DB 144 IPNYSFVRAHDNSQDLIAQIV 165

RESULT 5

US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 69.1%; Score 76; DB 10; Length 1781;
Best Local Similarity 59.1%; Pred. No. 0.0018;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSVQDLIRNII 22

DB 1123 IPNYSFVRAHDNSQDLIAQIV 1144


```
RESULT 6
US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CTP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match          64.5%; Score 71; DB 10; Length 584;
Best Local Similarity 63.2%; Pred. No. 0.0034;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYFIRAHSEVDLIR 19
Db      167 IPNYSFVRAHDYDADP 185

RESULT 7
US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CTP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match          61.8%; Score 68; DB 10; Length 545;
Best Local Similarity 70.0%; Pred. No. 0.0098;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 SYSFIRAHSEVDLIR 22
Db      156 NYIFIRAHSEVQTVIAK 175

RESULT 8
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Naga Prasad, Sathyamangla V.

RESULT 9
US-09-557-796-20
; Sequence 20, Application US/09557796
; Publication No. US20030073140A1
; GENERAL INFORMATION:
; APPLICANT: Hocht, James
; APPLICANT: Dattois, Veronique
; TITLE OF INVENTION: METABOLIC SELECTION METHODS
; FILE REFERENCE: 234/191
; CURRENT APPLICATION NUMBER: US/09/557,796
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/172,952
; PRIOR FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 282
; TYPE: PRT
; ORGANISM: YiaJ-EC
US-09-557-796-20

Query Match          40.0%; Score 44; DB 11; Length 282;
Best Local Similarity 43.5%; Pred. No. 40;
Matches 10; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY      2 PSY--SFRHSEVDLIR 22
Db      169 PDYKSYWESHCHIQPLTRNT 191

RESULT 10
US-10-101-235A-8
; Sequence 8, Application US/10101235A
; Publication No. US20030182669A1
; GENERAL INFORMATION:
; APPLICANT: Rockman, Howard A.
; APPLICANT: Naga Prasad, Sathyamangla V.

RESULT 6
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match          61.8%; Score 68; DB 9; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.029;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 SYSFIRAHSEVDLIR 22
Db      576 NYIFIRAHSEVQTVIAK 595
```

```

? APPLICANT: Laporte, Stephane A.
? APPLICANT: Barak, Larry S.
? APPLICANT: Caron, Marc G.
? TITLE OF INVENTION: Phosphoinositide 3-Kin
? FILE REFERENCE: 033072-064
? CURRENT APPLICATION NUMBER: US/10/101,235A
? CURRENT FILING DATE: 2002-03-19
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 8
? LENGTH: 1068
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-101-235A-8

```

Query Match	40.0%;	Score 44;	DB 12;	Length 1068;
Best Local Similarity	45.5%;	Pred. No. 1.8e+02;		
Matches 10;	Conservative 2;	Mismatches 10;	Indels 0;	Gaps 0;

QY 1 VPSYSFIRAHDSVEVQDLIRNII 22
:| | | | | | | | | | :
Db 123 MPVCEFDVMKDEVEQDFRRNII 144

RESULT 11
US-10-012-542-277
; Sequence 277, Application US/10012542
; Publication No. US20030044851A1
Copyright © 2003 by Inventor

```

1  APPLICANT: Ruben et al.
2  TITLE OF INVENTION: 94 Human Secreted Proteins
3  FILE REFERENCE: P202991
4  CURRENT APPLICATION NUMBER: US/10/012,542
5  CURRENT FILING DATE: 2001-12-12
6  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,355
7  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
8  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
9  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
10 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
11 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
12 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
13 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
14 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
15 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
16 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
17 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
18 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
19 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22

```

```

; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 277
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (340)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (341)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OS-10-012-542-2777

```

Query Match Similarity	39.1%	Score 43;	DB 15;	Length 347;
Best Local Similarity	35.3%	Pred. NO.	74;	
Matches	6;	Conservative	7;	Mismatches 4; Indels 0; Gaps 0

```
QY      4 YSFIRAHDSVQDLIRN 20
        : : : : : : : :
Db     103 HKYMRSHKAESELVRN 119
```

```

RESULT 12
US-10-156-761-11791
; Sequence 11791, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

```

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHITUKU
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11791

ORGANISM: Streptomyces avermitilis
US-10-156-761-11791

Query Match	39.1%	Score 43	DB 15	length 566
Best Local Similarity	47.4%	Pred. No.	1.3e+02	
Matches	9	Conservative	3	Mismatches 7
				Indels 0
				Gaps 0

```
QY      2 PSYSFIRAHDSVQDLIRN 20
         ||| : | : |||
Db      456 PSYRFLGPVDTAVGELTRN 474
```

US-10-177-293-470
; Sequence 470, Application US/10177293
; Publication No. US20030124128A1
GENERAL INFORMATION.

GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Glatz, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Ganmavaru, Manjula
APPLICANT: Kamatkar, Shubhang
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: East Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Puztari, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US 10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002

```

1  PRIOR FILING DATE: 2001-09-25
2  PRIOR APPLICATION NUMBER: US 60/362,585
3  PRIOR FILING DATE: 2002-03-05
4  PRIOR APPLICATION NUMBER: US 60/xxx,xxx
5  PRIOR FILING DATE: 2002-05-14
6  NUMBER OF SEQ ID NOS: 506
7  SOFTWARE: FASTSEQ for Windows Version 4.0.
8  SEQ ID NO 470
9  LENGTH: 2541
10 TYPE: PRT
11 ORGANISM: Homo sapiens
12 US-10-177-293-470

```

Query Match	39.1%;	Score 43;	DB 15;	Length 2541;
Best Local Similarity	38.1%;	Pred. No. 7.2e+02;		
Matches	8;	Conservative	4;	Mismatches 9;
				Indels 0;
				Gaps 0;

```
QY      2 PSYSFIRAHADSEVQDLIRNII 22
      ||| : | | | : : |
Db      1628 PSWSVLAGHSRTVSDSIKKLI 1648
```

RESULT 14
US-09-815-242-5358
; Sequence 5358, Application US/09815242
; Patent No. US20020061569A1
; Patent No. US20020061569A1

APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyeckind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes

```

1 CURRENT APPLICATION NUMBER: US/09/815,242
2 CURRENT FILING DATE: 2001-03-21
3 PRIOR APPLICATION NUMBER: 60/191,078
4 PRIOR FILING DATE: 2000-03-21
5 PRIOR APPLICATION NUMBER: 60/206,848
6 PRIOR FILING DATE: 2000-05-23
7 PRIOR APPLICATION NUMBER: 60/207,727
8 PRIOR FILING DATE: 2000-05-26
9 PRIOR APPLICATION NUMBER: 60/242,578
10 PRIOR FILING DATE: 2000-10-23
11 PRIOR APPLICATION NUMBER: 60/253,625
12 PRIOR FILING DATE: 2000-11-27
13 PRIOR APPLICATION NUMBER: 60/257,931
14 PRIOR FILING DATE: 2000-12-22
15 PRIOR APPLICATION NUMBER: 60/269,308
16 PRIOR FILING DATE: 2001-02-16
17 NUMBER OF SEQ ID NOS: 14110
18 SOFTWARE: FastSeq for Windows Version 4.0
19 SEQ ID NO 538
20 LENGTH: 1213
21 TYPE: PRT
22 ORGANISM: Staphylococcus aureus
23 US-09-815-242-538

```

Query Match	37.7%;	Score 41.5;	DB 9;	Length 1213;
Best Local Similarity	43.5%;	Pred. No. 5.4e+02;		
Matches 10; Conservative	6;	Mismatches 2;	Indels 5;	Gaps 1;

```
QY      5 SFIRAHDESEVOD-----LIRNIT 22
        ||:||:|:|:|:|:|:|:|:|:|
Db      615 SFLRAIDNPLQIDIXLVGLMRSVI 637
```

RESULT 15
US-09-815-242-12523

```

1 Sequence 12523, Application US/09815242
2 Patent No. US20020061569A1
3
4 GENERAL INFORMATION:
5 APPLICANT: Haselbeck, Robert
6 APPLICANT: Ohlsen, Karl L.
7 APPLICANT: Zyskind, Judith W.
8 APPLICANT: Wall, Daniel
9 APPLICANT: Trawick, John D.
10 APPLICANT: Carr, Grant J.
11 APPLICANT: Yamamoto, Robert T.
12 APPLICANT: Xu, H. Howard
13 TITLE OF INVENTION: Identification of Essential Genes in
14 TITLE OF INVENTION: Prokaryotes
15 FILE REFERENCE: EilatA.01A
16 CURRENT APPLICATION NUMBER: US/09/815,242
17 CURRENT FILING DATE: 2001-03-21
18 PRIOR APPLICATION NUMBER: 60/191,078
19 PRIOR FILING DATE: 2000-03-21
20 PRIOR APPLICATION NUMBER: 60/206,848
21 PRIOR FILING DATE: 2000-05-23
22 PRIOR APPLICATION NUMBER: 60/207,727
23 PRIOR FILING DATE: 2000-05-26
24 PRIOR APPLICATION NUMBER: 60/242,578
25 PRIOR FILING DATE: 2000-10-23
26 PRIOR APPLICATION NUMBER: 60/253,625
27 PRIOR FILING DATE: 2000-11-27
28 PRIOR APPLICATION NUMBER: 60/257,931
29 PRIOR FILING DATE: 2000-12-22
30 PRIOR APPLICATION NUMBER: 60/269,308
31 PRIOR FILING DATE: 2001-02-16
32
33 NUMBER OF SEQ ID NOS: 14110
34 SOFTWARE: FASTSEQ for Windows Version 4.0
35 SEQ ID NO 12523
36 LENGTH: 1217
37 TYPE: PRT
38 ORGANISM: Staphylococcus aureus
39
40 US-09-815-242-12523

```

```

Query Match      37.7% Score 41.5; DB 9; length 1217;
Best Local Similarity 43.5% Pred. No. 5.5e+02;
Matches 10; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

QY      5 SFRADSEVOD-----LIRNII 22
      |||::|::|::|
Db      615 SFRADINPELODIYLVGIMRSVI 637

```

Search completed: November 13, 2003, 10:29:04
Job time : 28.7773 secs

Page 5

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 14.2844 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-16
Perfect score: 107
Sequence: 1 MANYIFIRAHDSVQTVIAKII 22

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.dep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.dep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.dep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.dep:*
5: /cgn2_6/ptodata/2/iaa/PCUS.COMB.dep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.dep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	545	4	US-09-604-957-4
2	107	100.0	1430	3	US-09-008-172-2
3	107	100.0	1430	3	US-09-210-361-6
4	107	100.0	1430	4	US-09-740-274-6
5	89	83.2	1577	2	US-08-793-824-2
6	86	80.4	521	3	US-09-604-957-5
7	72	67.3	1475	3	US-09-007-999-2
8	72	67.3	1475	3	US-09-210-361-2
9	72	67.3	1475	3	US-09-740-274-2
10	68	63.6	1375	3	US-09-210-361-4
11	68	63.6	1375	4	US-09-740-274-4
12	55	51.4	584	4	US-09-604-957-6
13	53	49.5	2057	4	US-09-499-203-2
14	50	46.7	535	4	US-09-604-957-7
15	50	46.7	1278	4	US-09-604-957-3
16	44	41.1	183	4	US-09-107-532A-5511
17	40.5	37.9	279	4	US-09-107-532A-6807
18	40	37.4	349	4	US-09-585-858-57
19	40	37.4	871	4	US-09-585-858-36
20	40	37.4	2270	3	US-09-581-909-3
21	39	36.4	182	3	US-09-338-907-133
22	39	36.4	182	4	US-09-218-207-133
23	39	36.4	185	3	US-09-338-907-136
24	39	36.4	185	4	US-09-218-207-136
25	39	36.4	228	3	US-09-338-907-70
26	39	36.4	228	4	US-09-218-207-70
27	39	36.4	300	3	US-09-338-907-135

28	39	36.4	300	4	US-09-218-207-135	Sequence 135, App
29	39	36.4	315	3	US-09-338-907-134	Sequence 134, App
30	39	36.4	315	4	US-09-218-207-134	Sequence 134, App
31	39	36.4	353	2	US-08-996-306-4	Sequence 4, Appli
32	39	36.4	353	3	US-09-338-907-4	Sequence 4, Appli
33	39	36.4	353	4	US-09-218-207-4	Sequence 4, Appli
34	39	36.4	364	2	US-08-996-306-5	Sequence 5, Appli
35	39	36.4	364	3	US-09-338-907-5	Sequence 5, Appli
36	39	36.4	364	4	US-09-218-207-5	Sequence 5, Appli
37	38	35.5	121	4	US-09-134-001C-3077	Sequence 3077, Ap
38	38	35.5	146	2	US-08-647-960-8	Sequence 8, Appli
39	38	35.5	257	3	US-09-287-097-2	Sequence 2, Appli
40	38	35.5	260	2	US-08-685-992-4	Sequence 4, Appli
41	38	35.5	260	3	US-08-685-992-5	Sequence 4, Appli
42	38	35.5	260	2	US-09-144-925-4	Sequence 25, Appli
43	38	35.5	260	2	US-09-144-925-25	Sequence 25, Appli
44	38	35.5	261	2	US-08-685-992-3	Sequence 3, Appli
45	38	35.5	261	2	US-09-144-925-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-604-957-4
Sequence 4, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERITDINA HENDRIKA
APPLICANT: DIJKHITZEN, LUBBERT
APPLICANT: RAHAOUTI, HAKIM
APPLICANT: LEBER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 545
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match 100.0%; Score 107; DB 4; Length 545;
Best Local Similarity 100.0%; Pred. No. 4.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MANYIFIRAHDSVQTVIAKII 22
Db 154 MANYIFIRAHDSVQTVIAKII 175
RESULT 2
US-09-008-172-2
Sequence 2, Application US/09008172
Patent No. 6127602
GENERAL INFORMATION:
APPLICANT: Nicholas, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: Lataxes in Paper Manufacture
FILE REFERENCE: 0358D
CURRENT APPLICATION NUMBER: US/09/008,172
CURRENT FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1430
TYPE: PRT

```

; ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match      100.0%; Score 107; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  MANYIFIRAHDSVQTVIAKII 22
      |||||||
Db      574  MANYIFIRAHDSVQTVIAKII 595

RESULT 3
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match      100.0%; Score 107; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  MANYIFIRAHDSVQTVIAKII 22
      |||||||
Db      574  MANYIFIRAHDSVQTVIAKII 595

RESULT 4
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172

```

```

PRIORITY FILING DATE: 1998-01-16
PRIORITY APPLICATION NUMBER: 08/482,711
PRIORITY FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 6
LENGTH: 1430
TYPE: PROT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match      100.0%; Score 107; DB 4; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY      1 MANYFIRAHDSVQTVIAKII 22
      |||||||
Db      574 MANYFIRAHDSVQTVIAKII 595

RESULT 5
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
TITLE OF INVENTION: Genetic Manipulation of Plants to
NUMBER OF INVENTIONS: Increase Stored Carbohydrates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Griffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match      83.2%; Score 89; DB 2; Length 1577;
Best Local Similarity 81.0%; Pred. No. 2.7e-07;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY      2 ANYFIRAHDSVQTVIAKII 22
      |||||||
Db      660 ANYFIRAHDSVQTVIAKII 680

```

RESULT 6
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOU, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 4338
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 80.4%; Score 86; DB 4; Length 523;
Best Local Similarity 80.0%; Pred. No. 2.5e-07;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 148 NYSFIRADSEVQTVIAQIV 167

RESULT 7
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 67.3%; Score 72; DB 3; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.00029;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 554 SYSFIRADSEVQDLIADI 573

RESULT 8
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361

CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match 67.3%; Score 72; DB 3; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.00029;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 554 SYSFIRADSEVQDLIADI 573

RESULT 9
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 67.3%; Score 72; DB 4; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.00029;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 554 SYSFIRADSEVQDLIADI 573

RESULT 10
US-09-210-361-4

```
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4
```

```
Query Match      63.6%; Score 68; DB 3; Length 1375;
Best Local Similarity 70.0%; Pred. No. 0.0014;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      3 NYIFRAHDSVQTVIAKII 22
      |||:|||||:|:|:|
Db      580 SYSFIRAHDSVQDLIRNII 599
```

```
RESULT 11
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4
```

```
Query Match      63.6%; Score 68; DB 4; Length 1375;
Best Local Similarity 70.0%; Pred. No. 0.0014;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      3 NYIFRAHDSVQTVIAKII 22
      |||:|||||:|:|:|
Db      580 SYSFIRAHDSVQDLIRNII 599
```

```
RESULT 12
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LIEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL, GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 4338
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6
```

```
Query Match      51.4%; Score 55; DB 4; Length 584;
Best Local Similarity 55.6%; Pred. No. 0.11;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 NYIFRAHDSVQTVIAK 20
      |||:|||||:|:|:|
Db      169 NYSFVRAHDYDADPIRK 186
```

```
RESULT 13
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANH, Maria
; APPLICANT: KUETH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alernansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2
```

```
Query Match      49.5%; Score 53; DB 4; Length 2057;
Best Local Similarity 55.6%; Pred. No. 1.2;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      3 NYIFRAHDSVQTVIAK 20
      |||:|||||:|:|:|
Db      759 NYSFVRAHDYDADPIRK 776
```

```
RESULT 14
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
```

```

; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

```

Query Match          46.7%; Score 50; DB 4; Length 535;
Best Local Similarity 50.0%; Pred. No. 0.81;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

```

QY      3 NYIFRAHDSVQTVIAKII 22
Db      146 NYSFVRAHDNNSQDQIONAI 165

```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri.
US-09-604-957-3

```

```

Query Match          46.7%; Score 50; DB 4; Length 1278;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

```

QY      3 NYIFRAHDSVQTVIAKII 22
Db      622 NYSFVRAHDNNSQDQIONAI 641

```

Search completed: November 13, 2003, 09:48:02
 Job time : 15.2844 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 28.7773 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-16

Perfect score: 107

Sequence: 1 MANYIFIRAHDSHVQTVIAKII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	545	10	US-09-995-749A-10
2	107	100.0	1430	9	US-09-740-274-6
3	86	80.4	522	9	US-09-995-749A-11
4	72	67.3	1475	9	US-09-740-274-2
5	68	63.6	1375	9	US-09-740-274-4
6	53	49.5	584	10	US-09-995-749A-12
7	50	46.7	535	10	US-09-995-749A-13
8	46	46.7	1781	10	US-09-995-749A-2
9	44	41.1	583	10	US-09-758-269-2
10	42	39.3	327	12	US-10-032-585-7827
11	40	37.4	67	15	US-10-007-280A-214
12	40	37.4	349	15	US-10-270-875-57
13	40	37.4	349	15	US-10-270-878-57
14	40	37.4	349	15	US-10-270-786-57
15	40	37.4	349	15	US-10-270-710-57

16	40	37.4	349	15	US-10-270-859-57	Sequence 57, Appl
17	40	37.4	349	16	US-10-270-846-57	Sequence 56, Appl
18	40	37.4	871	15	US-10-270-875-36	Sequence 36, Appl
19	40	37.4	871	15	US-10-270-878-36	Sequence 36, Appl
20	40	37.4	871	15	US-10-270-786-36	Sequence 36, Appl
21	40	37.4	871	15	US-10-270-710-36	Sequence 36, Appl
22	40	37.4	871	15	US-10-270-859-36	Sequence 36, Appl
23	40	37.4	871	16	US-10-270-846-36	Sequence 36, Appl
24	39	36.4	78	9	US-09-864-761-41675	Sequence 41675, A
25	39	36.4	115	11	US-09-764-891-5098	Sequence 5098, Ap
26	39	36.4	182	10	US-09-901-489A-133	Sequence 133, App
27	39	36.4	182	10	US-09-853-526-133	Sequence 133, App
28	39	36.4	185	10	US-09-901-489A-136	Sequence 136, App
29	39	36.4	185	10	US-09-853-526-136	Sequence 136, App
30	39	36.4	228	10	US-09-901-489A-70	Sequence 70, Appl
31	39	36.4	228	10	US-09-853-526-70	Sequence 70, Appl
32	39	36.4	273	9	US-09-815-242-11100	Sequence 11100, A
33	39	36.4	300	10	US-09-901-489A-135	Sequence 135, App
34	39	36.4	300	10	US-09-853-526-135	Sequence 135, App
35	39	36.4	315	10	US-09-901-489A-134	Sequence 134, App
36	39	36.4	315	10	US-09-853-526-134	Sequence 134, App
37	39	36.4	353	10	US-09-901-489A-4	Sequence 4, Appl
38	39	36.4	353	10	US-09-915-582-59	Sequence 59, Appl
39	39	36.4	353	10	US-09-915-582-79	Sequence 79, Appl
40	39	36.4	353	10	US-09-853-526-4	Sequence 4, Appl
41	39	36.4	353	12	US-10-277-802-59	Sequence 59, Appl
42	39	36.4	353	12	US-10-277-802-79	Sequence 79, Appl
43	39	36.4	364	9	US-09-817-910-7	Sequence 7, Appl
44	39	36.4	364	10	US-09-901-489A-5	Sequence 5, Appl
45	39	36.4	364	10	US-09-853-526-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-995-749A-10

Sequence 10, Application US/09995749A

Patent No. US2002015568A1

GENERAL INFORMATION:

APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA HENDRIKA

APPLICANT: DIJKHUIZEN, LOUBERT

APPLICANT: RAHAOUI, HAKIM

APPLICANT: LEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

FILE REFERENCE: B043388-CIP

CURRENT APPLICATION NUMBER: US/09/995,749A

CURRENT FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: 09/604,957

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: EPO 00201871.1

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 545

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-995-749A-10

Query Match

Best Local Similarity 100.0%; Score 107; DB 10; Length 545;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 MANYIFIRAHDSHVQTVIAKII 22

154 MANYIFIRAHDSHVQTVIAKII 175

Db

US-09-740-274-6

Sequence 6, Application US/09740274

Patent No. US20020031826A1

```
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-6
```

```
Query Match          100.0%; Score 107; DB 9; Length 1430;
Best Local Similarity 100.0%; Pred. No. 4,8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MANYFIRAHDESEVQTVIAKII 22
DB 574 MANYFIRAHDESEVQTVIAKII 595
```

```
RESULT 3
; US-09-995-749A-11
; Sequence 11, Application US/09995749A
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LIMBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-VAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-995-749A-11
```

```
Query Match          80.4%; Score 86; DB 10; Length 522;
Best Local Similarity 80.0%; Pred. No. 4,9e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 NYIFIRAHDESEVQTVIAKII 22
DB 148 NYIFIRAHDESEVQTVIAKII 167
```

```
RESULT 4
; US-09-740-274-2
```

```
; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-2
```

```
Query Match          67.3%; Score 72; DB 9; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.0035;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 NYIFIRAHDESEVQTVIAKII 22
DB 554 SYSTFIRAHDESEVQTVIAKII 573
```

```
RESULT 5
; US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-4
```

```
Query Match          63.6%; Score 68; DB 9; Length 1375;
Best Local Similarity 70.0%; Pred. No. 0.015;
```

Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 NYIFRAHDESEVQTVIAKII 22
:|||||:|:|
Db 580 SYSFVRAHDNSQDQIQNAI 599

RESULT 6
US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOU, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 49.5%; Score 53; DB 10; Length 584;
Best Local Similarity 55.6%; Pred. No. 1.8;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 NYIFRAHDESEVQTVIAK 20
:|||||:|:|
Db 169 NYSFVRAHDYDQDPFRK 186

RESULT 7
US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOU, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 46.7%; Score 50; DB 10; Length 535;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 NYIFRAHDESEVQTVIAKII 22
:|||||:|:|
Db 146 NYSFVRAHDNSQDQIQNAI 165

RESULT 8
US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOU, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 46.7%; Score 50; DB 10; Length 1781;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 NYIFRAHDESEVQTVIAKII 22
:|||||:|:|
Db 1125 NYSFVRAHDNSQDQIQNAI 1144

RESULT 9
US-09-758-269-2
; Sequence 2, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSVERSE PLANTS CARRYING NEOXANTHIN
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-758-269-2

Query Match 41.1%; Score 44; DB 10; Length 583;
Best Local Similarity 47.4%; Pred. No. 58;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 YIFRAHDESEVQTVIAKII 22
:|||||:|:|
Db 530 YIFCHVADDETKTSEJQII 548

RESULT 10
US-10-032-585-7827
; Sequence 7827, Application US/10032585

```

RESULT 14 786-57
/ Sequence 57, Application US/10270786
/ Publication No. US20030087392A1
/ GENERAL INFORMATION:
/ APPLICANT: Sigridur Hjorleifsdatter
/ APPLICANT: Gudmundur O. Hrengyrdsson
/ APPLICANT: Olafur H. Fridjonsson
/ APPLICANT: Arnthor Aevartsson
/ APPLICANT: Jakob K. Kristjansson
/ TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
/ TITLE OF INVENTION: Host Organism
/ FILE REFERENCE: 2739.1001-001
/ CURRENT APPLICATION NUMBER: US/10/0270,786
/ CURRENT FILING DATE: 2002-10-11
/ PRIOR APPLICATION NUMBER: US/09/565,858
/ PRIOR FILING DATE: 2000-12-18

```

Query Match	37.4%;	Score 40;	DB 15;	Length 349;
Best Local Similarity	36.8%;	Pred. No. 1.5e+02;		
Matches	7;	Conservative	5;	Mismatches 7;
				Indels

QY 4 YIFIRAHDSVQGVIAKTI 22
| : | : | : | : | : | :
Db 175 YLIRSHIHEVNVISDKLL 193

```

RESULT 15
US-10-270-710-57
; Sequence 57, Application US/10270710
; Publication No. US20030092128A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevansson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739, 1001-001
; CURRENT APPLICATION NUMBER: US/10/270,710
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/565,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 349
; TYPE: PRF
; ORGANISM: Bacteriophage RM378
US-10-270-710-57

```

	Query March	37.4%	Score 40;	DB 15;	Length 349;
	Best local Similarity	36.8%	Pred. NO.	1.5e+02;	
Matches	7; Conservative	5;	Mismatches	7;	Indels 0;
Gaps					0;
Dy	4 YIFIRADSEFQVIYAKII	22			
	:: :: :::	:::			
Db	175 YLIRSRHIEENVVTSIDKLL	193			

Search completed: November 13, 2003, 10:29:05
Job time : 29.7773 secs

PI	Nichols SE;	
XX	WP1; 2002-414332/44.	
DR	N-PSDB; ABR52940.	
XX		
PT	Glucosyltransferase B or D protein useful for producing a glucan useful	
PT	as substitutes for and additions to modified starch and latexes in	
PT	paper manufacture, comprises mutations in specific positions	
XX		
PS	Disclosure; Page 38-42; 44pp; English.	
XX		
XX	The invention an isolated protein comprising a glucosyltransferase	
CC	(GTF) B polypeptide having changes at position from I448V, D457N,	
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,	
CC	D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,	
CC	Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having	
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and	
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,	
CC	an isolated polynucleotide which encodes P1 or P2, or its complementary	
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant	
CC	an expression cassette comprising the polynucleotide operably linked to a	
CC	promoter, a vector comprising the expression cassette, host cell	
CC	introduced with the vector, a transgenic plant comprising the	
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or	
CC	coating composition comprising a glucan produced in a plant, a latex,	
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex,	
CC	thermoplastic molecule or their combinations or glucan and starch where	
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line	
CC	deficient in starch biosynthesis, transformed with a gene encoding a	
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper	
CC	comprising the glucan (paper sizing/coating agent). The vector is useful	
CC	for producing a glucan in a plant. The method comprises transforming a	
CC	plant cell with the vector, growing the plant cell under plant growing	
CC	conditions to produce a regenerated plant and inducing expression of the	
CC	polynucleotide for a time sufficient to produce the glucan in the	
CC	regenerated plant, where the vector contains a transit sequence from	
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and	
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan	
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.	
CC	Glucans are useful as substitutes for and additions to modified starch	
CC	and latexes in paper manufacture. Unlike prior art techniques, which	
CC	require input materials that produce chemical effluents, paper	
CC	manufacture utilizing the glucan produced by GTF, which utilizes	
CC	biologically produced input materials, is more cost-effective and	
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic	
CC	properties and impart gloss to the paper during coating step.	
CC	The present sequence represents GTFD.	
XX		
XX	Sequence 1430 AA;	
DB		
DB	Query Match 100.0%; Score 107; DB 23; Length 1430;	
	Best Local Similarity 100.0%; Pred. No. 4.7e-10;	
	Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DB	1 MANTIFIRADSEYQTVIAKII 22	
	574 MANTIFIRADSEYQTVIAKII 595	
DB		
XX		
XX	AAU98043;	
XX		
DT	27-AUG-2002 (first entry)	
XX		
DE	S. mutans glucosyltransferase GTFD mutant N471D.	
XX		
KW	Glucosyltransferase; GTFD; transgenic plant; paper sizing;	
KW	coating composition; glucan; starch; latex; thermoplastic molecule;	
KW	amyloplast; vacuole; paper manufacture; mutant; muten.	
XX		
OS	Streptococcus mutans.	
XX	Synthetic.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 471	
XX	/note= "Wild-type Asn substituted by Asp"	
XX		
XX	US2002031826-A1.	
XX		
XX	14-MAR-2002.	
XX		
PF	19-DEC-2000, 2000US-0740274.	
XX		
XX	11-DEC-1998, 98US-0210361.	
PR	07-JUN-1995, 95US-0478704.	
PR	07-JUN-1995, 95US-0482711.	
PR	07-JUN-1995, 95US-0485243.	
PR	16-JAN-1998, 98US-0007999.	
PR	16-JAN-1998, 98US-0008172.	
PR	20-JAN-1998, 98US-0009620.	
XX		
XX	(NICH/) NICHOLS S E.	
XX		
XX	Nichols SE;	
XX	WP1; 2002-414332/44.	
XX		
PT	Glucosyltransferase B or D protein useful for producing a glucan useful	
PT	as substitutes for and additions to modified starch and latexes in	
PT	paper manufacture, comprises mutations in specific positions	
XX		
XX	Claim 36; Page -; 44pp; English.	
CC		
CC	The invention an isolated protein comprising a glucosyltransferase	
CC	(GTF) B polypeptide having changes at position from I448V, D457N,	
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,	
CC	D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,	
CC	Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having	
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and	
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,	
CC	an isolated polynucleotide which encodes P1 or P2, or its complementary	
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,	
CC	an expression cassette comprising the polynucleotide operably linked to	
CC	a promoter, a vector comprising the expression cassette, host cell	
CC	introduced with the vector, a transgenic plant comprising the	
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or	
CC	coating composition comprising a glucan produced in a plant transformed	
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex,	
CC	thermoplastic molecule or their combinations or glucan and starch where	
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line	
CC	deficient in starch biosynthesis, transformed with a gene encoding a	
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper	
CC	comprising the glucan (paper sizing/coating agent). The vector is useful	

Sequence 1430 AA;
Query Match 100.0%; Score 107; DB 23; Length 1430;
Best Local Similarity 100.0%; Pred. No. 4,7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANYIFIRAHSEVQVIKII 22
DB 574 MANYIFIRAHSEVQVIKII 595
RESULT 3
AAU98041
ID AAU98041 standard; Protein; 1430 AA.
AC AAU98041;
DT 27-AUG-2002 (first entry)
XX
XX S. mutans glucosyltransferase GTFD mutant T589D.
XX
XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
OS Streptococcus mutans.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 589 /note= "Wild-type Thr substituted by Asp"
FT
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX WPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex.

thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFD mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFD sequence appearing as AAU98029
CC and the information in claim 36.
XX
XX
XX Sequence 1430 AA;
SQ
Query Match 94.4%; Score 101; DB 23; Length 1430;
Best Local Similarity 95.5%; Pred. No. 5,7e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MANYIFIRAHSEVQVIKII 22
DB 574 MANYIFIRAHSEVQVIKII 595
RESULT 4
AAU98042
ID AAU98042 standard; Protein; 1430 AA.
XX
XX AAU98042;
XX
XX 27-AUG-2002 (first entry)
XX
XX
XX S. mutans glucosyltransferase GTFD mutant T589E.
XX
XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 589 /note= "Wild-type Thr substituted by Glu"
FT
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX

PI Nichols SE;
 XX
 DR WPI: 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -: 44pp; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTF sequence appearing as AAU98029
 CC and the information in claim 36.
 XX
 SQ Sequence 1430 AA;
 Query Match 94.4%; Score 101; DB 23; Length 1430;
 Best Local Similarity 95.5%; Pred. No. 5.7e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MANYIFIRAHDSVQVIAKII 22
 DB 574 MANYIFIRAHDSVQVIAKII 595
 RESULT 5
 AAU98044
 ID AAU98044 standard; Protein; 1430 AA.
 XX
 AC AAU98044;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFD mutant N471D/T589D.
 XX
 KM Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 471
 FT /note= "Wild-type Asn substituted by Asp"
 FT Misc-difference 589
 FT /note= "Wild-type Thr substituted by Asp"
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JUN-1998; 98US-0007999.
 PR 16-JUN-1998; 98US-0008172.
 PR 20-JUN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI: 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -: 44pp; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.

XX WPI: 1996-151376/15.
 DR N-PSDB; AATT13139.
 XX
 PT Plants contg. new bacterial DNA encoding glucosyl transferase
 PT activity - retain higher levels of stored carbohydrate(s) in a form
 PT readily digestible by ruminants
 XX
 PS Claim 4; Page 16-20; 31pp; English.
 XX
 CC The sequence represents an alpha-D-glucosyltransferase from
 CC Streptococcus salivarius. The enzyme is primer-independent, and
 CC produces soluble glucan from sucrose. A gene encoding the enzyme
 CC may be cloned and expressed in Escherichia coli using a subclone
 CC of phage lambda-C13, e.g. plasmid pSG501 or plasmid pSG502. The
 CC DNA may also be expressed in a transgenic plant, to improve the
 CC level of stored carbohydrate in a pasture plant which normally
 CC contains low levels, or to prevent degradation of stored carbohydrate
 CC during plant senescence. Dextran may be isolated from the plant, for
 CC use as a food binder or pharmaceutical additive. Primer independence
 CC ensures that the enzyme will be functional in plants. The glucan is
 CC poorly degraded in plants but easily degraded by bacteria in the rumen
 CC of grazing livestock.
 XX
 SQ Sequence 1577 AA;
 Query Match 83.2%; Score 89; DB 17; Length 1577;
 Best Local Similarity 81.0%; Pred. No. 9.4e-07;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ANYIFRADSEYQVYIAKII 22
 DB 660 ANYIFRADSEYQVYIAKII 680
 RESULT 8
 ABB98574
 ID ABB98574 standard; protein; 2835 AA.
 AC ABB98574;
 XX
 DT 14-JAN-2003 (first entry)
 XX
 DB Dextran saccharase, DSRE.
 XX
 KM Dextran saccharase; enzyme; cytotstatic; dermatological; antieborrheic;
 KM DSR-E; glycosyl transferase; dextran; prebiotic; pharmaceutical;
 KM microflora regulation; intestinal transit; mineral assimilation;
 KM colon cancer; acne; dandruff; body odour.
 XX
 OS Leuconostoc mesenteroides NRRL B-1299.
 XX
 FH Key
 FT Peptide
 FT Location/Qualifiers
 FT 1..40
 FT /label= Signal_peptide
 FT 41..2835
 FT /label= Mature_protein
 FT 423..439
 FT /label= SEQ_ID_6
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 478..501
 FT /label= SEQ_ID_7
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 519..539
 FT /label= SEQ_ID_8
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 560..571
 FT /label= SEQ_ID_9
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 631..645
 FT /label= SEQ_ID_10
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 1014..1021
 FT Peptide

FT /label= SEQ_ID_11
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2120..2138
 FT /label= SEQ_ID_12
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2161..2184
 FT /label= SEQ_ID_13
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2202..2214
 FT /label= SEQ_ID_14
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2243..2250
 FT /label= SEQ_ID_15
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2315..2322
 FT /label= SEQ_ID_16
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2689..2696
 FT /label= SEQ_ID_17
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 1981..1142
 FT /note= "Catalytic domain, SEQ ID 1. This sequence is
 FT specifically claimed in Claim 4"
 XX
 PN FR2022163-A1.
 XX
 PD 20-SEP-2002.
 XX
 PF 19-DEC-2001; 2001FR-0016495.
 XX
 PR 16-MAR-2001; 2001FR-0003631.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Bozonnet SAM, Renaud SMMC, Willemot RML, Monsan PEF.
 XX
 DR WPI: 2002-715213/78.
 DR N-PSDB; ABQ80961, ABQ80962.
 XX
 PT New glycosyl transferase enzymes, containing glucan bonding and
 PT catalytic domains and producing alpha-(1-2) branched dextran, useful
 PT in probiotic, pharmaceutical or cosmetic compositions
 XX
 PS Claim 6; Page 65-74; 82pp; French.
 XX
 CC The present sequence is a novel dextran saccharase, DSR-E, from
 CC Leuconostoc mesenteroides NRRL B-1299. The saccharase has glycosyl
 CC transferase activity suitable for producing dextran having alpha(1-2)
 CC branches from sucrose, p-nitrophenyl-alpha-D-glucopyranoside,
 CC alpha-fluoroglucose, alpha-D-glucopyranoside-alpha-D-sorbidopyranoside or
 CC alpha-D-galactopyranosyl-sucrose. The dextran saccharase is useful in
 CC prebiotic, pharmaceutical or cosmetic compositions. The dextran and
 CC related compounds having alpha(1-2) bonds, produced using DSR-E, may be
 CC involved in signalling/cellular recognition processes in vivo
 CC (specifically in regulation of microflora in the intestines or on the
 CC skin), and are potentially useful for improving intestinal transit,
 CC increasing assimilation of minerals (e.g. calcium and/or magnesium),
 CC preventing cancer of the colon and combating skin problems such as acne,
 CC dandruff and body odour.
 XX
 SQ Sequence 2835 AA;
 Query Match 83.2%; Score 89; DB 23; Length 2835;
 Best Local Similarity 90.0%; Pred. No. 1.9e-06;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 NYIFRADSEYQVYIAKII 22
 DB 630 NYIFRADSEYQVYIAKII 649
 RESULT 9
 AAU80055

ID AAU90055 standard; Protein; 1527 AA.
 AC AAU90055;
 DT 30-JUL-2002 (first entry)
 OS Leuconostoc mesenteroides dextranucrase.
 DE Dextranucrase; yoghurt; curd; cheese; fermented milk; infant formulae;
 KM pet food; vitamin; oral vaccine; enzyme.
 XX Leuconostoc mesenteroides.
 OS
 PN EPI201131-A1.
 PD 02-MAY-2002.
 PF 23-OCT-2000; 2000EP-0123012.
 XX 23-OCT-2000; 2000EP-0123012.
 PR 23-OCT-2000; 2000EP-0123012.
 XX (NEST) SOC PROD NESTLE SA.
 PA
 XX Bauche A, De Maleprade D, Duboc P, Neubauer H, Zink R;
 PI MPI: 2002-373873/41.
 DR N-PSDB; ARK50932.
 XX
 PS Disclosure; Page 14-20; 38pp; English.
 XX
 CC The invention relates to preparation of a fermented food product that
 CC includes treating the microorganism (A) to be used for fermentation with
 CC a dextranucrase so that this binds to the cell walls of (A). The
 CC method is used to produce yoghurt, curd, cheese or other fermented milk
 CC products, ice cream, fermented cereal products, water-based jellies,
 CC infant formulae and pet foods. Dextranucrase can also be bound to e.g.
 CC vitamins or oral vaccines for delivering these to foods, using (A) as
 CC carrier. The method produces foods with attractive and uniform texture,
 CC since dextranucrase is evenly distributed and forms a thickening agent
 CC in situ, with the amount formed being controlled by the amount of
 CC dextranucrase bound. Milk does not have to be coagulated in order to
 CC achieve a good texture, so the amount of acidic by-products formed is
 CC reduced. Dextranucrase binds to many different types of cells over wide
 CC ranges of temperature and pH and the treated bacteria can be stored in
 CC liquid or powdered form. The present sequence represents the amino acid
 CC sequence of dextranucrase.
 CC
 XX
 SQ Sequence 1527 AA;
 Query Match 80.4%; Score 86; DB 23; Length 1527;
 Best Local Similarity 80.0%; Pred. No. 3.1e-06;
 Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CY 3 NYIFRAHDSVQVTIAKII 22
 |||:|||||:|||||:
 Db 654 NYSEFRAHDSVQVTIAQIV 673
 RESULT 10
 AAU98036
 ID AAU98036 standard; Protein; 1475 AA.
 AC AAU98036;
 DT 27-AUG-2002 (first entry)
 OS S. mutans glucosyltransferase GTFB mutant D567T/D571K.
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;

KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 XX Synthetic.
 FT Key Location/Qualifiers
 FT Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 XX US2002031826-A1.
 XX
 XX
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH//) NICHOLS S E.
 XX
 PI Nichols SR;
 XX
 DR MPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 PS
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes p1 or p2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilising the glucan produced by GTF, which utilises
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 78.5%; Score 84; DB 23; Length 1475;
Best Local Similarity 85.0%; Pred. No. 6.9e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 NYIFRAHDSVQTVIAKII 22
DB 554 SYSFIRAHDSVQTVIAKII 573
RESULT 11
AAU98037
ID AAU98037 standard; Protein, 1475 AA.
XX
AC AAU98037;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
XX
KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; muten.
XX
OS Streptococcus mutans.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
FT
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX WPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -: 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T,
CC Y169R/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 78.5%; Score 84; DB 23; Length 1475;
Best Local Similarity 85.0%; Pred. No. 6.9e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 NYIFRAHDSVQTVIAKII 22
DB 554 SYSFIRAHDSVQTVIAKII 573
RESULT 12
AAU98038
ID AAU98038 standard; Protein, 1475 AA.
XX
XX AAU98038;
XX
XX 27-AUG-2002 (first entry)
XX
XX S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; muten.
XX
XX Streptococcus mutans.
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"
FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 779 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference /note= "Wild-type Lys substituted by Gln"

FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
 XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX an isolated polynucleotide which encodes P1 or P2, or its complementary
 XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX an expression cassette comprising the polynucleotide operably linked to a
 XX promoter, a vector comprising the expression cassette, host cell
 XX introduced with the vector, a transgenic plant comprising the
 XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 XX coating composition comprising a glucan produced in a plant transformed
 XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
 XX thermoplastic molecule or their combinations or glucan and starch where
 XX the glucan is produced in the amyloplast and/or vacuole or a maize line
 XX deficient in starch biosynthesis, transformed with a gene encoding a
 XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 XX comprising the glucan (paper sizing/coating agent). The vector is useful
 XX for producing a glucan in a plant. The method comprises transforming a
 XX plant cell with the vector, growing the plant cell under plant growing
 XX conditions to produce a regenerated plant and inducing expression of the
 XX polynucleotide for a time sufficient to produce the glucan in the
 XX regenerated plant, where the vector contains a transit sequence from
 XX ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 XX chlorophyll AB binding protein to produce a transgenic plant, and glucan
 XX is produced in the amyloplast of potato or the vacuole of sugar beet.
 XX Glucans are useful as substitutes for and additions to modified starch
 XX and latexes in paper manufacture. Unlike prior art techniques, which
 XX require input materials that produce chemical effluents, paper
 XX manufacture utilising the glucan produced by GTF, which utilises
 XX biologically produced input materials, is more cost-effective and
 XX environmentally friendly. Moreover, glucans also exhibit thermoplastic
 XX properties and impart gloss to the paper during coating step.
 XX The present sequence represents a GTFB mutant of the invention.
 XX Note: The present sequence is not shown in the specification but
 XX was created by the indexer using the GTFB sequence appearing as AAU98027
 XX and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 XX
 XX Query Match 78.5%; Score 84; DB 23; Length 1475;
 XX Best Local Similarity 85.0%; Pred. No. 6.9e-06;

Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 NYIFIRAHSEVQTVIAKII 22
 : |||||
 Db 554 SYIFIRAHSEVQTVIAKII 573
 RESULT 13
 AAU98032
 ID AAU98032 standard; Protein; 1475 AA.
 XX
 XX AAU98032;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB mutant D567T.
 XX
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 XX Streptococcus mutans.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 XX
 XX Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
 XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX an isolated polynucleotide which encodes P1 or P2, or its complementary
 XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX an expression cassette comprising the polynucleotide operably linked to a
 XX promoter, a vector comprising the expression cassette, host cell
 XX introduced with the vector, a transgenic plant comprising the
 XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 XX coating composition comprising a glucan produced in a plant transformed
 XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
 XX thermoplastic molecule or their combinations or glucan and starch where
 XX the glucan is produced in the amyloplast and/or vacuole or a maize line
 XX deficient in starch biosynthesis, transformed with a gene encoding a
 XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll a/b binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

XX Sequence 1475 AA;
XX
XX Query Match 72.9%; Score 78; DB 23; Length 1475;
XX Best Local Similarity 80.0%; Pred. No. 8.4e-05;
XX Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX 3 NTFFIRAHDSVQVIADII 22
XX :|||||:|||||:|||||:
XX 554 SYSFIRAHDSVQVIADII 573

RESULT 14
AAU98034
ID AAU98034 standard; Protein; 1475 AA.
XX
XX AAU98034;

XX 27-AUG-2002 (first entry)
XX
XX 5. mutans glucosyltransferase GTFB mutant D457N/D567T.
XX

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutlein.
XX

OS Streptococcus mutans.
OS Synthetic.
OS XX

XX Key Location/Qualifiers
XX

XX Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
XX FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
XX

XX US2002031826-A1.
XX

XX 14-MAR-2002.
XX

XX 19-DEC-2000; 2000US-0740274.
XX

XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX

XX (NICH/) NICHOLS S E.
XX

XX Nichols SE;
XX

DR WPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
XX as substitutes for and additions to modified starch and latexes in
XX paper manufacture, comprises mutations in specific positions
XX
XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
XX (GTF) B polypeptide having changes at position from I448V, D457N,
XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T,
XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
XX an isolated polynucleotide which encodes P1 or P2, or its complementary
XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
XX an expression cassette comprising the polynucleotide operably linked to a
XX promoter, a vector comprising the expression cassette, host cell
XX introduced with the vector, a transgenic plant comprising the
XX vector, a seed or tuber from the transgenic plant, a plant transformed
XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
XX thermoplastic molecule or their combinations or glucan and starch where
XX the glucan is produced in the amyloplast and/or vacuole or a maize line
XX deficient in starch biosynthesis, transformed with a gene encoding a
XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper
XX comprising the glucan (paper sizing/coating agent). The vector is useful
XX for producing a glucan in a plant. The method comprises transforming a
XX plant cell with the vector, growing the plant cell under plant growing
XX conditions to produce a regenerated plant and inducing expression of the
XX polynucleotide for a time sufficient to produce the glucan in the
XX regenerated plant, where the vector contains a transit sequence from
XX ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
XX chlorophyll a/b binding protein to produce a transgenic plant, and glucan
XX is produced in the amyloplast of potato or the vacuole of sugar beet.
XX Glucans are useful as substitutes for and additions to modified starch
XX and latexes in paper manufacture. Unlike prior art techniques, which
XX require input materials that produce chemical effluents, paper
XX manufacture utilizing the glucan produced by GTF, which utilizes
XX biologically produced input materials, is more cost-effective and
XX environmentally friendly. Moreover, glucans also exhibit thermoplastic
XX properties and impart gloss to the paper during coating step.
XX The present sequence represents a GTFB mutant of the invention.
XX Note: The present sequence is not shown in the specification but
XX was created by the indexer using the GTFB sequence appearing as AAU98027
XX and the information in claim 36.

XX Sequence 1475 AA;
XX

XX Query Match 72.9%; Score 78; DB 23; Length 1475;
XX Best Local Similarity 80.0%; Pred. No. 8.4e-05;
XX Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX

XX 3 NTFFIRAHDSVQVIADII 22
XX :|||||:|||||:|||||:
XX 554 SYSFIRAHDSVQVIADII 573

RESULT 15
AAU98035
ID AAU98035 standard; Protein; 1475 AA.
XX
XX AAU98035;

XX 27-AUG-2002 (first entry)
XX

XX 5. mutans glucosyltransferase GTFB mutant D457N/D571K.
XX

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutlein.
XX

OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 XX
 FN US2002031826-A1.
 PD
 PD 14-MAR-2002.
 XX
 PP 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 72.9%; Score 78; DB 23; Length 1475;
 Best Local Similarity 80.0%; Pred. NO. 8.4e-05;
 Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 NYIFIRAHDSVQTVIAKII 22
 : ||||| : |||||
 Db 554 SYSFIRAHDSVQDLIAKII 573

Search completed: November 13, 2003, 09:38:29
 Job time : 41.3507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6588 Seconds

(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-16

Sequence: 1 MANYIFIRAHDSVQTVIAKII 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	1431	2 A45866	dextranucrase (EC
2	90	84.1	1365	2 A41483	glucosyltransferas
3	89	83.2	1577	2 T30858	glucosyltransferas
4	88	82.2	1449	2 T30857	glucosyltransferas
5	88	82.2	1449	2 T30852	glucosyltransferas
6	87	81.3	1508	2 T31098	probable dextran
7	87	81.3	1599	2 S22737	glucosyltransferas
8	85	79.4	1290	2 JCS473	dextranucrase (EC
9	84	72.9	1518	2 A44811	glucosyltransferas
10	78	67.3	1475	2 B31335	gtfB protein precu
11	68	63.6	1375	2 UT0345	dextranucrase (EC
12	64	59.8	1592	2 A8175	glucosyltransferas
13	45	42.1	267	2 B84213	hypothetical prote
14	44.5	41.6	593	2 S51946	pyruvate kinase (E
15	44	41.1	132	2 D82080	mutator Mult prote
16	44	41.1	431	2 S49821	PR22 protein - Ara
17	44	41.1	583	2 T04531	nine-cis-epoxycar
18	44	41.1	1213	2 T04531	probable helicase
19	44	40.2	288	2 T26383	hypothetical prote
20	43	40.2	933	2 G70166	phosphate zinc prot
21	42	39.3	236	2 E82593	nitric-oxide regulon
22	42	39.3	1153	2 T31080	immunogenic protei
23	41.5	38.8	354	2 A70255	pyruvate kinase (E
24	41.5	38.8	493	2 T10054	pyruvate kinase (E
25	41.5	38.8	583	2 T10051	HCRF3 protein - hu
26	41	38.3	76	2 JQ1633	probable lipoprote
27	41	38.3	140	2 F70515	hypothetical prote
28	41	38.3	283	2 AF1276	uv excision repair
29	41	38.3	368	2 T40115	

30	41	38.3	411	2 A34526	ORF1 protein - Ory
31	41	38.3	411	2 T10373	hypothetical prote
32	41	38.3	602	2 A71256	aspartate-tRNA lig
33	41	38.3	725	2 T17732	helicase-like prot
34	41	38.3	768	2 S52684	probable membrane
35	41	38.3	1165	2 D72496	hypothetical prote
36	41	38.3	1302	2 AC2017	cobalamin biosynth
37	41	38.3	1771	2 S76851	hypothetical prote
38	40.5	37.9	217	1 H69797	rhinogalacturonan
39	40	37.4	344	1 GRYC87	sulfate transport
40	40	37.4	354	2 S65687	(A+T)-stretch-bind
41	40	37.4	382	2 G90540	hypothetical prote
42	40	37.4	417	2 S25627	glucose-1-phosphat
43	40	37.4	554	2 C69161	sensory transducti
44	40	37.4	557	2 S73434	aspartate-tRNA lig
45	40	37.4	586	2 B84271	glutamyl-tRNA synt

ALIGNMENTS

RESULT 1

A45866 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans

C/Species: Streptococcus mutans

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: A45866

R/Honda, O.; Kato, C.; Kuramitsu, H.K.

J. Gen. Microbiol. 136, 2099-2105, 1990

A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl

A/Reference number: A45866; MUID:91100958; PMID:2148600

A/Accession: A45866

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1431 <HON>

A/Cross-references: GB:M29296

C/Superfamily: cpl repeat homology

C/Keywords: glycosyltransferase; hexosyltransferase

F/181-201/Domain: cpl repeat homology <CP1>

F/1127-1146/Domain: cpl repeat homology <CP2>

F/1192-1211/Domain: cpl repeat homology <CP3>

F/1257-1276/Domain: cpl repeat homology <CP4>

F/1277-1297/Domain: cpl repeat homology <CP5>

F/1321-1340/Domain: cpl repeat homology <CP6>

F/1341-1361/Domain: cpl repeat homology <CP6>

F/1385-1404/Domain: cpl repeat homology <CP7>

Query Match 100.0%; Score 107; DB 2; Length 1431;

Best Local Similarity 100.0%; Pred. No. 2.5e-09; Mismatches 0; Gaps 0;

Matches 22; Conservative 0; Indels 0; Gaps 0;

QY 1 MANYIFIRAHDSVQTVIAKII 22

Db 574 MANYIFIRAHDSVQTVIAKII 595

RESULT 2

A41483 glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus

C/Species: Streptococcus sobrinus

C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999

C/Accession: A41483

R/Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

Infect. Immun. 58, 2452-2458, 1990

A/Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase

A/Reference number: A41483; MUID:90316665; PMID:2142479

A/Accession: A41483

A/Molecule type: DNA

A/Residues: 1-1365 <GIL>

A/Cross-references: GB:M30943; NID:G153652; PIDN:AAA26898.1; PID:G153653

C/Genetics:

A:Gene: gtfS

C/Superfamily: cpl repeat homology

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 84.1%; Score 90; DB 2; Length 1365;
Best Local Similarity 90.0%; Pred. No. 1.6e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRAHDSVQTVIAKII 22
DB 539 NYVFIKRAHDSVQTVIAKII 558

RESULT 3

T30858
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858

R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for p

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1577 <SIM>

A:Cross-references: EMBL:L35928; NID:G662380; PID:G662381; PIDN:AA041413.1

C:Genetics:

A:Gene: gtfm

Query Match 83.2%; Score 89; DB 2; Length 1577;
Best Local Similarity 81.0%; Pred. No. 2.8e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ANYFIRAHDSVQTVIAKII 22
DB 660 ANYFIRAHDSVQTVIAKII 680

RESULT 4

T30857
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30857

R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for p

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30857

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <SIM>

A:Cross-references: EMBL:L35495; NID:G662378; PID:G662379; PIDN:AA041412.1

C:Genetics:

A:Gene: gtfL

Query Match 82.2%; Score 88; DB 2; Length 1449;
Best Local Similarity 72.7%; Pred. No. 3.7e-06;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MANYFIRAHDSVQTVIAKII 22
DB 607 MANYFIRAHDSVQTVIAKII 628

RESULT 5

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30552

R:Jaaffe, R.I.

submitted to the EMBL Data Library, February 1998

A:Description: Streptococcus salivarius V1477 gtfN.

A:Reference number: Z20854

A:Accession: T30552

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <JAF>

A:Cross-references: EMBL:AF049609; NID:G2935545; PID:G2935546; PIDN:AA005156.1

C:Genetics:

A:Gene: gtfN

Query Match 82.2%; Score 88; DB 2; Length 1449;
Best Local Similarity 72.7%; Pred. No. 3.7e-06;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MANYFIRAHDSVQTVIAKII 22
DB 607 MANYFIRAHDSVQTVIAKII 628

RESULT 6

T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides

C:Species: Leuconostoc mesenteroides

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T31098

R:Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.

EMS Microbiol. Lett. 159, 307-315, 1998

A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS

A:Reference number: Z20981; MUID:98164374; PMID:9503626

A:Accession: T31098

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1508 <MON>

A:Cross-references: EMBL:AF030129; NID:G2766611; PID:G2766612; PIDN:AA095453.1

A:Experimental source: strain NRRL B-1299

C:Genetics:

A:Gene: dsrB

C:Function:

A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 81.3%; Score 87; DB 2; Length 1508;
Best Local Similarity 85.0%; Pred. No. 5.7e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRAHDSVQTVIAKII 22
DB 636 NYVFIKRAHDSVQTVIAKII 655

RESULT 7

S22737
glucosyltransferase (EC 2.4.1.1-) S - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000

C:Accession: S22737; S28810; B44811; S22727

R:Jacques, N. the EMBL Data Library, March 1992

submitted to the EMBL Data Library, March 1992

A:Reference number: S22726

A:Accession: S22737

A:Molecule type: DNA

A:Residues: 1-1599 <JAC>

A:Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAA77898.1; PID:G47531

A:Experimental source: ATCC 25975

R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A:Reference number: A44811; MUID:92148377; PMID:1838391

A:Accession: S28810

A:Molecule type: DNA

A:Residues: 1-51 <GTF>

A:Cross-references: EMBL:Z11873

C:Genetics:

A:Gene: gtfK
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 81.3%; Score 87; DB 2; Length 1599;
 Best Local Similarity 81.0%; Pred. No. 6.1e-06;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ANYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 573 ATYFIRAHDSVQTVIADII 593

RESULT 8
 JCS473
 dextranase (EC 2.4.1.5) - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C:Accession: JCS473
 R:Monchois, V.; Willemot, R.M.; Renaud-Simeon, M.; Croux, C.; Monsan, P.
 Gene 182, 23-32, 1996
 A:Title: Cloning and sequencing of a gene coding for a novel dextranase from Leuconostoc
 A:Reference number: JCS473; MUID:97136686; PMID:8982063
 A:Accession: JCS473
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1290 <MON>
 A:Cross-references: GB:U8181
 A:Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose onto
 A:Genetics:
 A:Gene: dsrA
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:78-870/Domain: catalytic #status predicted <CAT>
 F:922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 79.4%; Score 85; DB 2; Length 1290;
 Best Local Similarity 85.0%; Pred. No. 1e-05;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 NYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 389 NYSFIRAHDSVQTVIADII 408

RESULT 9
 A44811
 glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
 C:Species: Streptococcus salivarius
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
 C:Accession: A44811; S22726; S28809
 R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A:Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
 A:Reference number: A44811; MUID:92146377; PMID:1838391
 A:Accession: A44811
 A:Molecule type: DNA
 A:Residues: 1-1518 <GIF>
 A:Cross-references: EMBL:Z11873; NID:947526; PIDN:CAA77900.1; PID:947527
 A:Note: sequence extracted from NCBI backbone (NCBIN:81050, NCBIIP:81052)
 C:Genetics:
 A:Gene: gtfU
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 72.9%; Score 78; DB 2; Length 1518;
 Best Local Similarity 70.0%; Pred. No. 0.00018;
 Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 NYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 604 NYFIRAHDSVQTVIADII 623

RESULT 10
 B33135
 gtfB protein precursor - Streptococcus mutans

C:Species: Streptococcus mutans
 C>Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
 C:Accession: B33135; A33128
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: B33135
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1475 <SHI>
 A:Cross-references: GB:M17361; NID:9153639; PIDN:AAA88588.1; PID:9153640
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 submitted to the Protein Sequence Database, September 1990
 A:Reference number: A33128
 A:Accession: A33128
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-171,173-641, 'N', 643-1475 <SH2>
 A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 67.3%; Score 72; DB 2; Length 1475;
 Best Local Similarity 75.0%; Pred. No. 0.0018;
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 NYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 554 NYSFIRAHDSVQTVIADII 573

RESULT 11
 UT0345
 dextranase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
 N:Alternate names: sucrose 6-glucosyltransferase
 C:Species: Streptococcus mutans
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
 C:Accession: UT0345; C33135
 R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
 Gene 69, 101-109, 1988
 A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
 A:Reference number: UT0345; MUID:89137980; PMID:2976010
 A:Accession: UT0345
 A:Molecule type: DNA
 A:Residues: 1-1375 <UED>
 A:Experimental source: GS-5
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: C33135
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <SHI>
 A:Cross-references: GB:M17361
 C:Genetics:
 A:Gene: gtfC
 C:Function: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 A:Superfamily: cpl repeat homology
 C:Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-1375/Product: glucosyltransferase #status predicted <MAT>

F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1253-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 63.6%; Score 68; DB 2; Length 1375;
 Best Local Similarity 70.0%; Pred. No. 0.0078;
 Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 NYIFRADSEVQTVIAKII 22
 DB 580 SYSFARADSEVQDIIRNII 599

RESULT 12

A:8175
 C:Species: Streptococcus sobrinus
 C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999

A:Accession: A38175
 R:Abdo, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
 J. Bacteriol. 173, 989-996, 1991

A>Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
 A:Reference number: A38175; MUID:91123227; PMID:1704006

A:Accession: A38175
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-1592 <ABO>
 A:Cross-references: GB:D90213; NID:9217032; PIDN:BA014241.1; PID:dl014946; PID:g217033

C:Superfamily: cpl repeat homology
 F:1093-1112/Domain: cpl repeat homology <CP1>
 F:1222-1241/Domain: cpl repeat homology <CP2>
 F:1287-1306/Domain: cpl repeat homology <CP3>
 F:1330-1351/Domain: cpl repeat homology <CP4>
 F:1352-1371/Domain: cpl repeat homology <CP5>
 F:1402-1420/Domain: cpl repeat homology <CP6>
 F:1465-1484/Domain: cpl repeat homology <CP7>
 F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 59.8%; Score 64; DB 2; Length 1592;
 Best Local Similarity 65.0%; Pred. No. 0.043;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 NYIFRADSEVQTVIAKII 22
 DB 550 SYSFARADSEVQDIIRNII 569

RESULT 13

A:84213
 C:Species: Halobacterium sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

A:Accession: B84213
 R:Ng, W.Y.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Madocke, D.G.; Jablo
 Tung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A>Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: B84213
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-267 <STO>

A:Cross-references: GB:AE004437; NID:gi0580148; PIDN:AA019070.1; GSPDB:GN0C138
 A:Gene: eif2a

C:Superfamily: translation initiation factor eIF-2 alpha chain

Query Match 42.1%; Score 45; DB 2; Length 267;
 Best Local Similarity 50.0%; Pred. No. 8.8;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 IRADSEVQTVIAKII 22
 DB 52 VRDHVEDQTVIAKVL 67

RESULT 14

A:51946
 C:Species: Nicotiana tabacum (common tobacco)
 C>Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 18-Jun-1999

A:Accession: S51946; S44286
 R:Blakeley, S.; Gottlob-McHugh, S.; Wan, J.; Crews, L.; Miki, B.; Ko, K.; Dennis, D.T.
 Plant Mol. Biol. 27, 79-89, 1995

A>Title: Molecular characterization of plastid pyruvate kinase from castor and tobacco.
 A:Reference number: S51946; MUID:95170010; PMID:7865798

A:Accession: S51946
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-593 <BLA>

A:Cross-references: EMBL:Z28373
 R:Blakeley, S.; Gottlob-McHugh, S.; Wan, J.; Crews, L.; Miki, B.; Ko, K.; Dennis, D.

submitted to the EMBL Data Library, November 1993
 A:Description: Molecular characterization of plastid pyruvate kinase from castor and tob

A:Reference number: S44286
 A:Accession: S44286

A:Molecule type: mRNA
 A:Residues: 1-86, A', 88-593 <BLM>

A:Cross-references: EMBL:Z28373; NID:9482935; PIDN:CA062222.1; PID:9482935
 A:Gene: nuclear
 C:Superfamily: pyruvate kinase
 C:Keywords: chloroplast; glycolysis; phosphotransferase

Query Match 41.6%; Score 44.5; DB 2; Length 593;
 Best Local Similarity 52.4%; Pred. No. 26;
 Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 MANYIFRADSEVQTVIAKII 21
 DB 323 LKSYIQARSDI-SVIAKI 342

RESULT 15

A:82080
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

A:Accession: D82080
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, P.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Frazer, C.M.
 Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: D82080
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-132 <HEI>
 A:Cross-references: GB:AE004310; GB:AE003852; NID:9656963; PIDN:AA095335.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:

A:Gene: VC2392
 A:Map position: 1

C:Superfamily: mutator mutT; mutT domain homology

Query Match 41.1%; Score 44; DB 2; Length 132;
 Best Local Similarity 50.0%; Pred. No. 5.9;
 Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Thu Nov 13 12:12:17 2003

us-09-290-049a-16.rpt

Page 5

Search completed: November 13, 2003, 09:50:30
Job time : 15.6588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.32417 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-16

Perfect score: 107
Sequence: 1 MANYIFRAHDSVQTVIAKII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	1462	1	GTFD_STRMU
2	90	84.1	1365	1	GPR5_STRDO
3	68	63.6	1455	1	GTFD_STRMU
4	67	62.6	1476	1	GTFB_STRMU
5	64	59.8	1592	1	GTF2_STRDO
6	62	57.9	1597	1	GTF1_STRDO
7	45	42.1	267	1	IF2A_HAUNI
8	44.5	41.6	593	1	KRYA_TOBAC
9	44	41.1	479	1	PRL2_ARATH
10	42	39.3	205	1	ADEN_ADEB8
11	42	39.3	236	1	PHOU_XYLF
12	42	39.3	336	1	RA51_DROME
13	42	39.3	1153	1	NOS_LYMS
14	41.5	38.8	583	1	KRYA_RICCO
15	41	38.3	368	1	RH23_SCHPO
16	41	38.3	411	1	VP48_NVOOP
17	41	38.3	602	1	SYD_TREPA
18	40	37.4	344	1	CYS5_SYND7
19	40	37.4	417	1	AGP_PROBE
20	40	37.4	557	1	SYD_MYCPN
21	40	37.4	654	1	CBP1_YEAST
22	40	37.4	1018	1	VGNM_BPNV
23	40	37.4	1246	1	SKIM_HUMAN
24	39.5	36.9	231	1	BIOD_YEAST
25	39.5	36.9	332	1	COBS_PSEDE
26	39.5	36.9	1082	1	RRPO_RORPC
27	39.5	36.9	146	1	V224_FOMPY
28	39	36.4	238	1	Y381_TREBA
29	39	36.4	246	1	O8336_FEWLP
30	39	36.4	283	1	R12_HAELN
31	39	36.4	297	1	LEG1_HAECC
32	39	36.4	297	1	LEC3_CAEEL
33	39	36.4	350	1	VOD2_DROME

34	39	36.4	351	1	RFB_XANCP	P5525 xanthomonas
35	39	36.4	354	1	PLCE_HUMAN	Q9NUQ2 homo sapien
36	39	36.4	380	1	ADH_MALDO	P48977 malus domes
37	39	36.4	433	1	TCR_STANU	P02983 straphylococ
38	39	36.4	434	1	SYN_PYPUR	O8U43 pyrococcus
39	39	36.4	461	1	GTFB_FLEBA	O9FAX2 flexibacter
40	39	36.4	525	1	COX1_CAEEL	P24893 caenorhabdi
41	39	36.4	548	1	LIP2_CANRU	P32946 candida rug
42	39	36.4	549	1	LIP3_CANRU	P32947 candida rug
43	39	36.4	549	1	LIP4_CANRU	P32948 candida rug
44	39	36.4	586	1	SYD_EORPU	O51402 borrelia bu
45	39	36.4	592	1	SYD_THETN	Q8RA17 thermocaneer

ALIGNMENTS

RESULT 1	GTFD_STRMU	STANDARD; PRT; 1462 AA.
ID	GTFD_STRMU	AC P49331; 069383; 069386; 069389; 069392; 069398;
DT	01-FEB-1996 (Rel. 33, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)	
DE	(Sucrose 6-glucosyltransferase).	
GN	GTFD OR SMU.910.	
OS	Streptococcus mutans.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1309;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=GS-5;	
RC	MEDLINE=91100958; PubMed=2148600;	
RA	Honda O., Kato C., Kuramitsu H.K.;	
RT	"Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme.";	
RL	J. Gen. Microbiol. 136:2099-2105(1990).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F, MT4467 / Serotype B, and MT8148 / Serotype C;	
RC	MEDLINE=98231643; PubMed=9570124;	
RA	Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.;	
RA	Kimura S., Hamada S.;	
RT	"Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans.";	
RT	FEMS Microbiol. Lett. 161:331-336(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=UA159 / ATCC 700610 / Serotype C;	
RC	MEDLINE=2295063; PubMed=12397186;	
RA	Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.;	
RA	Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.;	
RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perrecci J.;	
RT	"Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).	
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.	
CC	-1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-fucose + ((1,6)-alpha-D-glucosyl) (N+1).	
CC	-1- SUBCELLULAR LOCATION: Secreted.	
CC	-1- DISEASE: DENTAL CARIES.	
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.	
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.	

```

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29236; AAA26895.1; -
DR EMBL; D88653; BAA26103.1; -
DR EMBL; D88656; BAA26107.1; -
DR EMBL; D88659; BAA26111.1; -
DR EMBL; D88662; BAA26115.1; -
DR EMBL; D89979; BAA26121.1; -
DR EMBL; AE014932; AAN58619.1; -
DR InterPro; IPR002479; CW_binding_70.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Complete proteome.
FT SIGNAL 1 1462
FT CHAIN 2 1462
FT DOMAIN 1232 1423
FT REPEAT 1232 1295
FT REPEAT 1296 1359
FT REPEAT 1360 1423
FT REPEAT 10 10
FT VARIANT 19 19
FT VARIANT 58 58
FT VARIANT 68 68
FT VARIANT 81 81
FT VARIANT 113 113
FT VARIANT 122 122
FT VARIANT 132 132
FT VARIANT 135 135
FT VARIANT 137 137
FT VARIANT 202 202
FT VARIANT 255 255
FT VARIANT 275 275
FT VARIANT 288 288
FT VARIANT 301 301
FT VARIANT 313 313
FT VARIANT 317 317
FT VARIANT 328 328
FT VARIANT 350 350
FT VARIANT 628 628
FT VARIANT 688 688
FT VARIANT 726 726
FT VARIANT 730 730
FT VARIANT 762 762
FT VARIANT 964 964
FT VARIANT 1019 1019
FT VARIANT 1059 1059
FT VARIANT 1060 1060
FT VARIANT 1080 1080
FT VARIANT 1142 1142
FT VARIANT 1198 1198
FT VARIANT 1220 1220
FT VARIANT 1280 1280
FT VARIANT 1282 1282
FT VARIANT 1290 1290
FT VARIANT 1311 1311
FT VARIANT 1311 1311

```

```

FT VARIANT 1403 1403 G -> D (IN STRAINS GS-5 AND MT4467).
FT VARIANT 1425 1425 G -> R (IN STRAIN GS-5).
FT VARIANT 1449 1449 R -> K (IN STRAIN MT4467).
FT CONFLICT 1428 1462 RYDKNSGMNYNKKVTLANRRIGIDRWGIARY -> VY
FT R (IN REF. 1)
SQ SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;
Query Match 100.0%; Score 107; DB 1; Length 1462;
Best Local Similarity 100.0%; Pred. No. 7.1e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANYFIPIAHDSQVTVIAKII 22
DB 574 MANYFIPIAHDSQVTVIAKII 595
RESULT 2
GTF5_STRDO STANDARD; PRT; 1365 AA.
ID GTF5_STRDO
AC P29336;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTF5.
OS Streptococcus downei (Streptococcus sobrius).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus;
OC NCB_TaxID=1317;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell R.R., Ferretti J.J.;
RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans.";
RL Infect. Immun. 58:2452-2458 (1990).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30943; AAA26898.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36
FT CHAIN 37 1365
FT REPEAT 157 177
FT REPEAT 178 197
FT DOMAIN 198 1061
FT REPEAT 1062 1082

```

```
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
SQ SEQUENCE 1365 AA, 151590 MW, 1672965A2E8C476 CRC64,

Query Match 84.1%, Score 90; DB 1; Length 1365;
Best Local Similarity 90.0%; Pred. No. 5e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 3 NYIFIRAHDSVQVIAKII 22
Db 539 NYVIFIRAHDSVQVIAKII 558

RESULT 3
GTFPC_STRMU STANDARD; PRT; 1455 AA.
AC P13470; O69382; O69388; O69391; O69397; P05427;
DT 01-NOV-1998 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update) (GTF-SI)
DE Glucosyltransferase-St precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextrantransferase) (Sucrose 6-glucosyltransferase).
GN GTFPC OR SMU.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=96231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar C., Lai H., White J., Roe B.A., Perretti J.U.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP SEQUENCE OF 1-349 FROM N.A.
RX STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
```



```
FT  VARIANT 1377 1377 R -> K (IN STRAIN MT8148).
FT  VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT  VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT  VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT  VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT  CONFLICT 1337 1455 ORLYKSNQVQAGKELLTERGRIKIVYPPNSGENVRNNYVR
TSQGMWYNGDVALIGMHVVEGRVRYFDNNGVYRASHD
GRNHWDYRRDPGRGSSSAVFRHSRNGFDNFRF ->
HASILSLMVFRLRESLSQSVAYVSNIMLIFEMKFIYIM
(IN REF. 1).
SQ  SEQUENCE 1455 AA; 162965 MW; 3CB455A994FEC86 CRC64;

Query Match 63.6%; Score 68; DB 1; Length 1455;
Beat Local Similarity 70.0%; Pred. No. 0.0029;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 NYIFIRAHSEVQVIYAKII 22
Db 580 SYSFIRAHSEVQDIIRNII 599

RESULT 4
GTFB STRMU STANDARD; PRT; 1476 AA.
AC P08987; 069381; 069387; 069390; 069396;
DT 01-NOV-1988 (Rel. 09; Created)
DT 28-SEP-2003 (Rel. 41; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
GN GTFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-GS-5;
RC MEDLINE=87308013; PubMed=3040685;
RA Shirota T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfb gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-UA159 / ATCC 700610 / Serotype C;
RC MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McClughlin R.E., Savic G., Chang J.,
RA Caron M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC
```

```

FT. VARIANT      1329 1329      H -> Y (IN STRAINS GS-5 AND MT4467).
FT. VARIANT      1394 1394      Y -> H (IN STRAINS GS-5, MT4239, MT4467
FT. VARIANT      1402 1402      AND MT8148).
FT. VARIANT      1459 1459      S -> G (IN STRAINS GS-5, MT4239, MT4467
FT. VARIANT      1459 1459      AND MT8148).
FT. CONFLICT     570 570      Y -> H (IN STRAIN MT4467).
FT. CONFLICT     800 817      R -> A (IN REF. 1).
FT. CONFLICT     800 817      ADDVRVAASTABSTDSK -> LIRKFALELAPHOOMA
SQ. SEQUENCE     1310 1310      (IN REF. 1).
SQ. SEQUENCE     1476 AA; 165846 MW; 9C6E09F731B4CBFC CRC64;

Query Match      62.6%; Score 67; DB 1; Length 1476;
Best Local Similarity 70.0%; Pred. No. 0.0043;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 NYIFIRAHSEVQVTIAKII 22
Db 554 SYSFIRAHSEVQDILRDII 573

RESULT 5
GTF2_STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; Pubmed=1704006;
RA Abo H.; Matsumura T.; Kodama T.; Ohba H.; Kato K.;
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase)";
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90213; BAAL4241.1; -
CC InterPro; IPR002479; CM binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CM binding_1; 13
CC Pfam; PF02324; Glyco_hydro_70; 1.

```

```

KM Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL      1 38      POTENTIAL.
FT CHAIN      39 1592      GLUCOSYLTRANSFERASE-I.
FT DOMAIN     1093 1592      CATALYTIC (APPROXIMATE).
FT DOMAIN     1093 1592      GLUCAN-BINDING (APPROXIMATE)..
FT DOMAIN     1093 1592      6.5 X TANDEM REPEATS.
FT REPEAT     1093 1142      1.
FT REPEAT     1158 1207      2.
FT REPEAT     1222 1272      3.
FT REPEAT     1287 1337      4.
FT REPEAT     1402 1451      5.
FT REPEAT     1514 1563      6.
FT REPEAT     1577 1592      7 (INCOMPLETE).
SQ. SEQUENCE     1592 AA; 176167 MW; BC0A66D079351BCF CRC64;

Query Match      59.8%; Score 64; DB 1; Length 1592;
Best Local Similarity 65.0%; Pred. No. 0.015;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 NYIFIRAHSEVQVTIAKII 22
Db 550 SYSFARAHSEVQDILRDII 569

RESULT 6
GTF1_STRDO STANDARD; PRT; 1597 AA.
AC ID GTF1_STRDO
AD P11001;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=87308014; Pubmed=3040686;
RA Ferretti J.J.; Gilpin M.L.; Russell R.R.B.;
RA "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus Mfe28.";
RL J. Bacteriol. 169:4271-4278(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17391; AAC63063.1; -
CC InterPro; IPR002479; CM binding.
CC InterPro; IPR003318; Glyco_hydro_70.

```



```

FT ACT SITE 341 341 BY SIMILARITY.
FT METAL 343 343 MAGNESIUM (POTENTIAL).
FT METAL 364 364 MAGNESIUM (POTENTIAL).
FT METAL 365 365 MAGNESIUM (POTENTIAL).
SQ SEQUENCE 593 AA; 65227 MW; 5AF049E193C1D484 CRC64;

Query Match 41.6%; Score 44.5; DB 1; Length 593;
Best Local Similarity 52.4%; Pred. No. 11;
Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Qy 1 MANYFIRADSEVQYTIK 21
Db 323 LKSYIQARASDI-SYIYAKI 342

RESULT 9
PRL2 ARATH STANDARD; PRT; 479 AA.
ID PRL2 ARATH 094007; Q9LUR9;
AC 039150; 094007; Q9LUR9;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PPI/PP2A phosphatases pleiotropic regulator PRL2.
OS PRL2 OR AT3G1650 OR MGL6.11.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=9843845; PubMed=9765207;
RA Nemeth K., Salchert K., Putnocky P., Bhalerao R., Koncz-Kalman Z.,
RA Stankovic-Stangeland B., Bako L., Mathur J., Okiesz L., Stabel S.,
RA Gebenberger P., Sitt M., Redei G.P., Scheil J., Koncz C.;
RT "pleiotropic control of glucose and hormone responses by PRL1, a
RT nuclear WD protein, in Arabidopsis."
RT Genes Dev. 12:3059-3073(1998).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and PAC
RT clones."
RT DNA Res. 7:131-135(2000).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (ANG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLEIOTROPIC REGULATOR OF PPI AND PP2A PHOSPHATASES.
CC -1- SIMILARITY: Contains 7 WD repeats.
CC -1- SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD-REPEAT PROTEINS.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X82826; CAA58033.1;
CC EMBL; AB022217; BAB02756.1; ALT SEQ.
CC EMBL; AY054181; AAL06842.1; ALT_INIT.

```

```

DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPRTENBRPT.
DR PRODOM; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KM Repeat, WD repeat.
FT REPEAT 168 198 WD 1.
FT REPEAT 210 240 WD 2.
FT REPEAT 252 282 WD 3.
FT REPEAT 293 323 WD 4.
FT REPEAT 335 364 WD 5.
FT REPEAT 377 406 WD 6.
FT REPEAT 426 456 WD 7.
SQ SEQUENCE 479 AA; 53568 MW; 3019864AAE80670 CRC64;

Query Match 41.1%; Score 44; DB 1; Length 479;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 5 IFRADSEVQYTIK 20
Db 289 IFVLPDSEVQYTIK 304

RESULT 10
ADEN_ADEG8 STANDARD; PRT; 205 AA.
ID ADEN_ADEG8 09QW72;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenin (EC 3.4.22.39) (Endoprotease) (Late L3 23 kDa protein).
OS Avian adenovirus type 8 (Strain ATCC A-24) (Fowl adenovirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Avadenovirus.
OX NCBI_TaxID=66295;
RN 1
RP SEQUENCE FROM N.A.
RA Ojick D., Nagy E.;
RT "The DNA sequence of fowl adenovirus 8."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIOLE PROTEINASE CLEAVING SPECIFIC GLY-ALA PEPTIDES IN A
CC NUMBER OF VIRAL PRECURSOR PROTEINS (TERMINAL, IIA, VI, VII, VIII,
CC IIA AND MU) WHICH ARE REQUIRED FOR VIRION MATURATION. ALSO CLEAVES
CC HOST CELLS CYTOSKELETAL KERATINS K7 AND K18.
CC -1- CATALYTIC ACTIVITY: Cleaves proteins of the adenovirus and its
CC host cell at two consensus sites: -Yaa-Xaa-Gly-Gly- and
CC -Yaa-Xaa-Gly-Xaa-|-Gly- (in which Yaa is Met, Ile or Leu, and Xaa
CC is any amino acid).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF083975; AAD50345.2;
CC HSRP; P03252; IAVP.
CC MEROPS; C05.001; -.
DR InterPro; IPR000855; Peptidase_C5.
DR Pfam; PF00770; Peptidase_C5; 1.
DR PRINTS; PR00703; ADENOPDASE.
DR PRODOM; PD003705; Peptidase_C5; 1.
DR Hydrolase; Thiol protease; Late protein.
KW ACT_SITE 55
KW ACT_SITE 72
KW ACT_SITE 122
SQ SEQUENCE 205 AA; 23701 MW; 36F0700CDFB85F62 CRC64;

```

Query Match 39.3%; Score 42; DB 1; Length 205;
 Best Local Similarity 53.3%; Pred. No. 9;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MANYIFIRADSEVOY 15
 DB 179 MKNSFPAHSEIK 193

RESULT 11

PHOU_XYLEFA STANDARD; PRT; 236 AA.
 AC Q9PB09 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphate transport system protein phou.
 GN PHOU OR XP2145.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 CC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=23711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.B., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Faccin A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Graber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Matsukuma A.Y.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Monteiro C.B.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Peguero J.B.,
 RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.U.M.,
 RA da Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.B.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zacc M., Mendes U., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-159(2000).
 CC -1- FUNCTION: NOT KNOWN; PROBABLY INVOLVED IN PHOSPHATE TRANSPORT
 CC AND/OR METABOLISM (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PHOU FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE004029; AAF84944.1; -;
 DR PIR; E82593; E82593.
 DR InterPro; IPR002754; Phou.
 DR Pfam; PF01895; Phou; 2. Complete proteome.
 KW Phosphate transport; Complete proteome.

SQ SEQUENCE 236 AA; 26596 MW; 2B0FB29DA2F14D1 CRC64;

Query Match 39.3%; Score 42; DB 1; Length 236;
 Best Local Similarity 30.0%; Pred. No. 10;
 Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVOYTIKII 22
 DB 6 NYHIVKSYDDENRLVTEII 25

RESULT 12

RA51_DROME STANDARD; PRT; 336 AA.
 AC Q27237; Q9VAA8;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA repair protein Rad51 homolog (RecA protein homolog).
 GN RAD51 OR DMR OR CG7948.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=95161094; PubMed=7857671;
 RA Akaboshi E., Inoue Y., Iryo H.;
 RT "Cloning of the cDNA and genomic DNA that correspond to the recA-like
 RL gene of Drosophila melanogaster."
 RL Jpn. J. Genet. 69:663-670(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=96207535; PubMed=8625736;
 RA McKee B.D., Ren X.T., Hong C.S.;
 RT "A recA-like gene in Drosophila melanogaster that is expressed at
 RT high levels in female but not male meiotic tissues."
 RL Chromosoma 104:479-488(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkelley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Andrews-Pfannkoch C., Baldwin D.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Basley E.M.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doull L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulyov G., Milshina N.V., Mobarry C., Morris J., Moshirefi A.,
 RA Mountrif S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusse D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Stylianias R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster." ;
RL Science 287:2185-2195(2000) .
CC -1- FUNCTION: BINDS TO SINGLE AND DOUBLE STRANDED DNA AND EXHIBITS
CC -1- DNA-DEPENDENT ATPASE ACTIVITY. UNDERWINDS DUPLICATION DNA (BY
CC SIMILARITY) .
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable) .
CC -1- SIMILARITY: BELONGS TO THE RECA FAMILY. RAD51 SUBFAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on res-
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-ch) .
-----
CC DR EMBL; D37788; BAA07039.1; -
CC DR EMBL; D17726; BAA04580.1; -
CC DR EMBL; LA1342; AAA64873.1; -
CC DR EMBL; AE003772; AAF57005.1; -
CC HSSP; O06609; 1822
CC FLYBase; FBgn0011700; Rad51.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR000445; HHH_1.
CC InterPro; IPR003583; HHH_1.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00633; HHH; 1.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00278; HHH; 1.
CC PROSITE; PS50162; RECA_2; 1.
CC PROSITE; PS50163; RECA_3; 1.
CC KW DNA-binding; ATP-binding; Nuclear protein.
CC NP BIND 124 131 ATP (POTENTIAL) .
CC SQ SEQUENCE 336 AA; 36647 MW; F9P9B21405B15DB0 CRC64;
-----
OY Query Match 39.3%; Score 42; DB 1; Length 336;
DB Best Local Similarity 42.1%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0.
-----
OY 3 NYIFRADSEVQVIYAKI 21
DB 185 NVAFIRANSDQDTKLQIM 203
-----
RESULT 13
NOS LYMSI
AC ID NOS LYMSI STANDARD; PRT; 1153 AA.
AC O61309;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nitric-oxide synthase (EC 1.14.13.39) (NOS, type I) (Neuronal NOS)
DE (N-NOS) (NOS) .
GN NOS.
OS Lymnaea stagnalis (Great pond snail) .
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Bascomatophora;
OC Lymnaeidae; Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RX [1]
RX SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT) .
RX TISSUE=CNS;
EX MEDLINE=9821186; PubMed=9552167;
EX Korneev S.A., Piper M.R., Picot J., Phillips R., Korneeva E.I.,
RT O'Shea M.;
RT "Molecular characterization of NOS in a mollusc: expression in a giant

```

```

modulatory neuron." ;
Rt J. Neurobiol. 35:65-76(1998) .
CC -1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY (BY SIMILARITY).
CC CC -1- CATALYTIC ACTIVITY: L-arginine + N NADPH + H+O(2) = citrulline + nitric oxide + N NADP(+).
CC CC -1- COFACTOR:HEME. BINDS ONE MOLE EACH OF FAD AND FMN (BY SIMILARITY).
CC CC -1- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Long;
IsoId=O61309-1; Sequence=Displayed;
CC CC Name=Short;
CC CC IsoId=O61309-2; Sequence=VSP_003584;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CENTRAL NERVOUS SYSTEM, IN THE SEROTONERGIC CEREBRAL GIANT CELLS. BOTH THE LONG AND SHORT ISOFORMS ARE EXPRESSED EQUALLY IN THE CNS.
CC -1- SIMILARITY: BELONGS TO THE NOS FAMILY.
CC -1- SIMILARITY: Contains 1 flavodoxin-like domain.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi-sib.ch).
CC -----
DR EMBL: AF012531; AAC17487.1; -.
DR PIR: T31080; T31080.
DR HSSP: P29477; 1DD7.
DR InterPro: IPR003097; FAD binding.
DR InterPro: IPR001094; Flavodoxin like.
DR InterPro: IPR001226; Flavodoxin.
DR InterPro: IPR001709; FPN CYT redctse.
DR InterPro: IPR004030; NO synthase.
DR InterPro: IPR001433; Oxid_FAD/NAD(P).
DR Pfam: PF00667; FAD_binding_1; 1.
DR Pfam: PF00258; flavodoxin; 1.
DR Pfam: PF00175; NAD_binding_1; 2.
DR Pfam: PF02898; NO synthase; 1.
DR PRINTS: PR00369; FLAVODOXIN.
DR PRINTS: PR00371; FPMCR.
DR PROSITE: PS05902; FLAVODOXIN LIKE; 1.
DR PROSITE: PS60001; NOS; 1.
DR KEGG: Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme; Repeat; Alternative splicing.
KW DOMAIN
FT 427 610 FLAVODOXIN-LIKE.
FT METAL 82 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT DOMAIN 397 417 CALMODULIN-BINDING (POTENTIAL).
FT NP_BIND 556 587 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 697 708 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 836 846 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 911 929 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 1089 1104 NADP (ADP PART) (BY SIMILARITY).
FT DOMAIN 934 1010 11 X 7 AA TANDEM REPEATS OF E-[NT]-[ST]-[IM]-P-S-C.
FT VASAPLIC 276 309 Missing (in isoform short).
FT /FTId=VSP_003584.
SO SEQUENCE 1153 AA; 129085 MW; 101877D02B6B109 CRC64;

Query Match 39.3%; Score 42; DB 1; Length 1153;
Best Local Similarity 38.1%; Pred. No. 57;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0.

QY 1 MANTIFRAHDSVCVTYATK 21
:|::||:|::|:
Db 667 LAERIQLAKSDDOETLIK 687

RESULT 14
KEYA_RICCO
```

ID KEVA_RICCO STANDARD; PRT; 583 AA.
 AC Q43117; Q43118;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyruvate Kinase Isozyme A, chloroplast precursor (EC 2.7.1.40).
 OS Ricinus communis (Castor bean).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucosids I; Malviales; Euphorbiaceae; Ricinus.
 CC NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA), AND SEQUENCE OF 75-81.
 RC STRAIN=cv. Baker 296; TISSUE=Endosperm;
 RA Blakeley S.D., Plaxton W.C., Dennis D.T.;
 RT "Relationship between the subunits of leucoplast pyruvate kinase from
 Ricinus communis and a comparison with the enzyme from other
 sources.";
 RL Plant Physiol. 96:1283-1288(1991).
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate = ADP + phosphoenolpyruvate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM (BY SIMILARITY).
 CC -1- PATHWAY: GLYCOLYSIS; final step.
 CC -1- SUBUNIT: OLIGOMER OF ALPHA AND BETA SUBUNITS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Alpha;
 CC IsoId=Q43117-1; Sequence=Displayed;
 CC Name=Beta;
 CC IsoId=Q43117-2; Sequence=VSP_002886;
 CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL, M64736; AAA33870.1; -
 CC DR EMBL, M64737; AAA33871.1; -
 CC DR PIR, T10051; T10051.
 CC DR PIR, T10054; T10054.
 CC DR HSSP, P14178; 1E0T.
 CC DR InterPro; IPR001697; Pyruvate_kinase.
 CC DR Pfam; PF00224; PK_1.
 CC DR Pfam; PF02887; PK_C_1.
 CC DR PRINTS; PR01050; PYRUVTKINASE.
 CC DR PRODOM; PD001009; Pyruvate_kinase; 1.
 CC DR TIGRFAMs; TIGR01064; PYRUV_Kin_1.
 CC DR PROSITE; PS00110; PYRUVATE_KINASE; 1.
 CC KW Pyruvate; Transferase; kinase; Glycolysis; Magnesium; Chloroplast;
 KW Transit peptide; Alternative splicing
 FT TRANSIT 1 74 CHLOROPLAST.
 FT CHAIN 75 583 PYRUVATE KINASE ISOZYME A.
 FT ACT SITE 331 331 BY SIMILARITY.
 FT METAL 333 333 MAGNESIUM (POTENTIAL).
 FT METAL 354 354 MAGNESIUM (POTENTIAL).
 FT METAL 355 355 MAGNESIUM (POTENTIAL).
 FT VARSPIC 1 129 MQQSIFSPNLTFRKQPEPKLPPPTNSRYPVNNYSLS
 FT IKAISTPSSSSDPQIVVAANGNGNSGVLYNNNKAVTSDP
 FT SLAEVAVTELEKENGFSRTKIVCTVCTGRTGFEELF
 FT ALAVGG -> MAVVYVDLEAVRVVAVLALRMEVVAVLT
 FT AVGVGD -> (in isoform Beta).
 FT FTID=VSP_002886.
 FT SEQUENCE 583 AA; 64093 MW; 5DFDE4F78A755C2 CRC64;
 Query Match 38.8%; Score 41.5; DB 1; Length 583;
 Best Local Similarity 47.6%; Pred. No. 33;
 Matches 10; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

DB 313 LKSYIAKRSDDI-AVIKI 332
 : : : : :
 RESULT 15
 RH23_SCHPO
 ID RH23_SCHPO STANDARD; PRT; 368 AA.
 AC 074803;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE UV excision repair protein rhp23 (RAD23 homolog).
 GN RHP23 OR SPEC2D10.12.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
 RC STRAIN=Sp223;
 RX MEDLINE=21648918; PubMed=11788722;
 RA Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,
 RA Zhao Y.;
 RT "Involvement of rhp23, a Schizosaccharomyces pombe homolog of the
 RT human HHR23A and Saccharomyces cerevisiae Rad23 nucleotide excision
 RT repair genes, in cell cycle control and protein ubiquitination.";
 RL Nucleic Acids Res. 30:581-591(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes D., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.D., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolckstein G., Aert R., Roben J., Grymoprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Filz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Rayon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovsky G.V., Ussery D., Bartell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [3]
 RP CHARACTERIZATION.
 RC MEDLINE=20119127; PubMed=10652237;
 RA Lombaerts M., Goeloe J.I., den Dulk H., Brandsma J.A., Brouwer J.;
 RT "Identification and characterization of the rhp23(+) DNA repair gene
 RT in Schizosaccharomyces pombe.";
 RL Biochem. Biophys. Res. Commun. 268:210-215(2000).
 CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
 CC postreplication repair functions in gap-filling of a daughter
 CC strand on replication of damaged DNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC -1- SIMILARITY: Contains 2 UBA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL: AF174293; AAD51975.1; -
DR EMBL: AL031788; CAA21170.1; -
DR PIR: T40115; T40115.
DR HSSP: P54725; IDV0.
DR GeneDB Spombe; SPBC2D10.12; -
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PSS0030; UBA; 2.
DR PROSITE; PSS0053; UBQUITIN 2; 1.
DR DNA damage; DNA repair; Nuclear protein; Repeat.
KM DOMAIN 1 77 UBQUITIN-LIKE.
FT DOMAIN 135 185 UBA 1.
FT DOMAIN 320 360 UBA 2.
FT DOMAIN 119 122 POLY-ALA.
FT DOMAIN 205 208 POLY-GLN.
SQ SEQUENCE 368 AA; 40135 MM; SCE75EB7E190EFPD CRC64;

Query Match 38.3%; Score 41; DB 1; Length 368;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ANYFIRAHDE 13
Db 354 ANYLFEHGHSE 365

Search completed: November 13, 2003, 09:45:34
Job time : 8.92417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 33.1564 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-16
Perfect score: 107

Sequence: 1 MANYIFRAHDSFVQTIAKII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP rivirus: *
16: SP bacteriophage: *
17: SP archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	93.5	1575	2	091CH3 streptococc
2	95	88.8	1577	2	054178 streptococc
3	90	84.1	1338	2	09WXJ4 streptococc
4	89	83.2	1577	2	055265 streptococc
5	89	83.2	2835	2	08G9Q2 leucococc
6	88	82.2	1449	2	068542 streptococc
7	88	82.2	1449	2	055264 streptococc
8	87	81.3	1016	2	091CJ7 leucococc
9	87	81.3	1477	2	091466 leucococc
10	87	81.3	1508	2	09EZH5 leucococc
11	87	81.3	1508	2	052224 leucococc
12	87	81.3	1512	2	09WXJ5 streptococc
13	87	81.3	1599	2	000599 streptococc
14	86	80.4	1527	2	09ZAR4 leucococc
15	86	80.4	1527	2	08KRE1 leucococc
16	85	79.4	1290	2	048756 leucococc

17	84	78.5	1554	2	08KZL5 streptococc
18	78	72.9	1518	2	000600 streptococc
19	64	59.8	1590	2	059983 streptococc
20	64	59.8	1590	2	055263 streptococc
21	53	49.5	2057	2	09RE05 leucococc
22	48.5	45.3	771	16	08EDZ6 shewanella
23	47	43.9	1197	5	09B121 drosophila
24	47	43.9	1197	5	09VCH8 drosophila
25	47	43.9	1387	5	0964N5 drosophila
26	46.5	43.5	783	5	08SX46 drosophila
27	46.5	43.5	866	5	09V8T7 drosophila
28	46	43.0	881	16	08G5W2 bifidobacte
29	45	42.1	368	16	08K610 streptococc
30	45	42.1	437	16	099YP6 streptococc
31	45	42.1	437	16	08P020 streptococc
32	45	42.1	533	10	08H6T2 chlamydomon
33	45	42.1	665	5	09VAL1 Q9V4J2 drosophila
34	45	42.1	805	16	08DDJ2 vibrio vuln
35	44	41.1	132	16	09KPH6 vibrio chol
36	44	41.1	264	16	08D618 vibrio chol
37	44	41.1	341	2	09L7T7 rhodospirill
38	44	41.1	583	10	049505 araldopsis
39	44	41.1	694	2	0915M3 salmonella
40	44	41.1	694	16	0935G9 salmonella
41	44	41.1	1213	3	059801 schistosach
42	43	40.2	261	17	P94945 methanopyru
43	43	40.2	288	5	09N8M7 caenorhabdi
44	43	40.2	933	16	051486 dorrellia bu
45	42	39.3	214	16	08ED17 shewanella

ALIGNMENTS

RESULT 1

091CH3 ID 091CH3 PRELIMINARY; PRT; 1575 AA.

AC 01-OCN-2000 (TREMBLrel. 15, Created)
DT 01-OCN-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCN-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557; PubMed=10768934;
RX MBLIN=20231779; Pubmed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.,
RT "Purification, characterization, and molecular analysis of the gene
encoding glucosyltransferase from Streptococcus oralis.";
RI Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BAA95201.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2B543 CRC64;

Query Match 93.5%; Score 100; DB 2; Length 1575;
Best Local Similarity 90.9%; Pred. No. 7.2e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MANYIFRAHDSFVQTIAKII 22

DB 617 MANYIFRAHDSFVQTIAKII 638

RESULT 2

```

Q54178      PRELIMINARY;      PRT; 1577 AA.
ID   Q54178; Q54247;
AC   Q54178; Q54247;
DT   01-NOV-1996 (TEMBLrel. 01, Created)
DT   01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE   Glucosyltransferase.
GN   GTFG.
OS   Streptococcus gordonii Challis.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=29390;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CHALLIS;
RA   MEDLINE=96157084; PubMed=8586195;
RA   Vickerman M.M.; Sulaevik M.C.; Clewell D.B.;
RT   "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT   phase variants";
RL   Dev. Biol. Stand. 85:309-314(1995).
RN   [2]
RP   SEQUENCE OF 1-96 FROM N.A.
RC   STRAIN=CHALLIS;
RC   MEDLINE=92276337; PubMed=1534326;
RA   Sulaevik M.C.; Tardif G.; Clewell D.B.;
RT   "Identification of a gene, rgg, which regulates expression of
RT   glucosyltransferase and influences the Spp phenotype of Streptococcus
RT   gordonii Challis.";
RL   J. Bacteriol. 174:3577-3586(1992).
DR   EMBL; U12643; AAC3483.1; -
DR   EMBL; M89776; AAA26969.1; -
DR   InterPro; IPR002479; CW binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 18.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
KW   Transferase.
SQ   SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match      88.8%; Score 95; DB 2; Length 1577;
Best Local Similarity 90.5%; Pred. No. 5.1e-07;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ANYIFRAHDSVQTVIAKII 22
DB      620 ANYIFRAHDSVQTVIAKII 640

RESULT 3
ID   Q9WKJ4      PRELIMINARY;      PRT; 1338 AA.
AC   Q9WKJ4;
DT   01-NOV-1999 (TEMBLrel. 12, Created)
DT   01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE   GTF-S.
GN   GTF-S.
OS   Streptococcus criceti.
OC   Plasmid pAK1.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1333;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HS-6;
RA   Inoue M.; Fukui K.; Miyagi A.;
RT   "S.cricetus glucosyltransferase(gtfS and gtfT) genes";
RT   Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AB026123; BAA77236.1; -
DR   InterPro; IPR002479; CW binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 10.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
KW   Plasmid.

```

```

SQ   SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D39B CRC64;

Query Match      84.1%; Score 90; DB 2; Length 1338;
Best Local Similarity 90.0%; Pred. No. 3e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 ANYIFRAHDSVQTVIAKII 22
DB      511 NYVIFRAHDSVQTVIAKII 530

RESULT 4
ID   Q55265      PRELIMINARY;      PRT; 1577 AA.
AC   Q55265;
DT   01-NOV-1996 (TEMBLrel. 01, Created)
DT   01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE   Glucosyltransferase precursor.
GN   GTFM.
OS   Streptococcus salivarius.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1304;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   MEDLINE=95122197; PubMed=7822030;
RA   Simpson C.L.; Giffard P.M.; Jacques N.A.;
RT   "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT   coding for primer-independent glucosyltransferases.";
RL   Infect. Immun. 63:609-621(1995).
DR   EMBL; L35928; AAC41413.1; -
DR   InterPro; IPR004828; Cutinase_antigen.
DR   InterPro; IPR002479; CW_binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 11.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
DR   ProDom; PD15343; Cutinase_antigen; 1.
KW   Signal; Transferase.
KW   SIGNAL.
FT   CHAIN      39..1577      POTENTIAL.
FT   CHAIN      39..1577      GLUCOSYLTRANSFERASE.
SQ   SEQUENCE 1577 AA; 175290 MW; 3EFP898A7D3A7BF3 CRC64;

Query Match      83.2%; Score 89; DB 2; Length 1577;
Best Local Similarity 81.0%; Pred. No. 5.4e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 ANYIFRAHDSVQTVIAKII 22
DB      660 ANYIFRAHDSVQTVIAKII 680

RESULT 5
ID   Q8G9Q2      PRELIMINARY;      PRT; 2835 AA.
AC   Q8G9Q2;
DT   01-MAR-2003 (TEMBLrel. 23, Created)
DT   01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE   Dextranucrase (EC 2.4.1.5) (Fragment).
GN   DSRE.
OS   Leuconostoc mesenteroides.
OC   Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX   NCBI_TaxID=1245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   MEDLINE=22231661; PubMed=12270834;
RA   Bozonnet S.; Dols-Lafargue M.; Fabre E.; Pizzut S.; Renaud-Simeon M.;
RA   Mondan P.; Willemot R.M.;
RT   "Molecular characterisation of DSR-E, an alpha-1,2 linkage
RT   synthesising dextranucrase with two catalytic domains.";
RL   J. Bacteriol. 184:5753-5761(2002).
DR   EMBL; AF430204; CAD22883.1; -

```

```

KW      Transferase; Glycosyltransferase
FT     NON TER          1
SQ     SEQUENCE       2835 AA;   313264 MW;   D03262CDD73539D CRC64;

Query Match
Best Local Similarity    83.2%; Score 89; DB 2; Length 2835;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      3 NYIFRAHDSVQTVIAKII 22
        ||| ||||| ||||| ||||| |||
Db       630 NYAFIRAHDSVQTVIAQII 649

RESULT 6
O68542      PRELIMINARY; PRT; 1449 AA.
ID O68542; AC O68542;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DR 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DN Glucosyltransferase N (Fragment).
GN GFEN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus
CX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 gtfN.";
RL Submitted (FEB-1998) to the EMBL/genbank/DDBJ databases.
DR EMBL; AF049609; AAC05156.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1, 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DK Transferrase.
KM NON TER
FT FT
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BFB CRC64;

Query Match
Best Local Similarity    82.2%; Score 88; DB 2; Length 1449;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0

OY      1 MANYIFIRAHDSVQTVIAKII 22
        ||| |:::||||| ::| :||
Db       607 MANYAFVRAHDSVQSIIQII 628

RESULT 7
O55264      PRELIMINARY; PRT; 1449 AA.
ID O55264; AC O55264;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DR 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DN Glucosyltransferase precursor.
GN GFPL.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus
CX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.W., Jacques N.A.;
RT coding for primer-independent glucosyltransferases";
RT Infect. Immun. 63:609-621(1995).
DR EMBL; U35495; AAC41412.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.

```

```

DR Pfam: PF01473; CW_binding_1; 8.
DR Pfam: PF02324; Glyco_hydro_70; 1.
FT SIGNAL; Transferase. 35
FT CHAIN 36 1449
SQ SEQUENCE 1449 AA; 159984 MW; D622F07306E86A46 CRC64;

Query Match 82.2%; Score 88; DB 2; Length 1449;
Best Local Similarity 72.7%; Pred. No. 7.2e-06;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 MANYFIRADSEVQTVIAKII 22
   |||:|||||:::|
Db 607 MANYAFVRAHDSVQITIGQII 628

RESULT 8
O9LCJ7 PRELIMINARY; PRT; 1016 AA.
AC O9LCJ7;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxId=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
RL mesenteroides NRRL B-512F.";
RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
DR EMBL; AB020020; BAA90527.1; -.
DR HSRP; P06278; 1A1S.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFD613CCB47 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1016;
Best Local Similarity 85.0%; Pred. No. 7.3e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 NYIFIRADSEVQTVIAKII 22
   |||:|||||:::|
Db 626 NYSFVRAHDSVQTVIAELI 645

RESULT 9
O9LA66 PRELIMINARY; PRT; 1477 AA.
ID O9LA66
AC O9LA66;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRC.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxId=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Remaud-Simeon M., Pizut S., Sarcabal P.,
RA Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternanucrase, a sucrose
RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.",
RL Submitted (OCT-1999) to the EMBL/GenBank/DBD databases.
RL EMBL; AJ250172; CAB75565.1; -.
DR InterPro; IPR002479; CW_binding.

```

DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEFCB831 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1477;
 Best Local Similarity 85.0%; Pred. No. 1.1e-05;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 605 NYSFVRAHDSVQTVIAQII 624

RESULT 10
 Q9EZH5 PRELIMINARY; PRT; 1508 AA.

AC Q9EZH5; 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Dextranucrase DsrB742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-742CB;
 RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF294469; AAG38021.1;
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87AEAF3A CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1508;
 Best Local Similarity 85.0%; Pred. No. 1.1e-05;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 636 NYSFVRAHDSVQTVIAQII 655

RESULT 11
 OS2224 PRELIMINARY; PRT; 1508 AA.

AC OS2224; 01-JUN-1998 (TRENBLREL. 06, Created)
 DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Renaud-Simeon M., Monan P., Willemot R.M.;
 RT "Cloning and sequencing of a gene coding for an extracellular
 dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
 RT synthesizing only a 4(1-6) glucan."
 RL EMBL; AF030129; AAB95453.1;
 DR EMBL; AF030129; AAB95453.1;
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.

DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 168511 MW; E70CBCE57A70D1F0 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1508;
 Best Local Similarity 85.0%; Pred. No. 1.1e-05;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 636 NYSFVRAHDSVQTVIAQII 655

RESULT 12
 Q9WKJ5 PRELIMINARY; PRT; 1512 AA.

AC Q9WKJ5; 01-NOV-1999 (TRENBLREL. 12, Created)
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE GTF.
 GN GTF.
 OS Streptococcus criceti.
 OC Streptococcus criceti.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OX NCBI_TaxID=1333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB026123; BAA77237.1;
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Plasmid.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1512;
 Best Local Similarity 80.0%; Pred. No. 1.1e-05;
 Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 561 SYFVRAHDSVQTVIAQII 580

RESULT 13
 Q00599 PRELIMINARY; PRT; 1599 AA.

AC Q00599; 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Glucosyltransferase S precursor (EC 2.4.1.5) (GTF) (Dextranucrase)
 GN (Sucrose 6-glucosyltransferase).
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=93381463; PubMed=8371114;
 RA Gifford P.M., Allen D.M., Milward C.P., Simpson C.L., Jacques N.A.;
 RT "Sequence of the gtf gene of Streptococcus salivarius ATCC 25975 and
 RT evolution of the gtf genes of oral streptococci."
 RL J. Gen. Microbiol. 139:1511-1522 (1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=92148377; PubMed=1838391;
 RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
 RT "Molecular characterization of a cluster of at least two
 RT glucosyltransferase genes in *Streptococcus salivarius* ATCC 25975.";
 RL J. Gen. Microbiol. 137:2577-2592 (1991).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
 CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
 CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-
 CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- SIMILARITY: TO REGIONS OF BARLEY AND BACILLUS AMYLOLIQUEFACIENS
 CC ALPHA AMYLASES AND RABBIT GLYCOGEN PHOSPHORYLASE.
 DR EMBL; Z11872; CAJ77898.1; -;
 DR EMBL; Z11873; CAJ77901.1; -;
 DR EMBL; M6411; AAA26897.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 42 POTENTIAL.
 FT CHAIN 43 1599 GLYCOSYLTRANSFERASE S.
 SQ SEQUENCE 1599 AA; 176480 MW; 24877869E152B707 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1599;
 Best Local Similarity 81.0%; Pred. No. 1.2e-05;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 NYIFIRAHSEVQTVIAKII 22
 DB 573 ATYIFVRAHSEVQTVIAQIV 593

RESULT 14
 Q9ZAR4 PRELIMINARY; PRT; 1527 AA.
 AC Q9ZAR4;
 DT 01-MAY-1999 (TRENBLREL. 10, Created)
 DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Dextranucrase.
 GN DEX.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512-F;
 RT "Cloning and Molecular Characterization of Dextranucrase Gene from
 RT Leuconostoc mesenteroides NRRL B-512F.";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U81374; AAD10352.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1527 AA; 169709 MW; 1DFAFA237C743398 CRC64;

Query Match 80.4%; Score 86; DB 2; Length 1527;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRAHSEVQTVIAKII 22
 DB 654 NYSFVRAHSEVQTVIAQIV 673

RESULT 15
 Q9KRE1 PRELIMINARY; PRT; 1527 AA.
 AC Q9KRE1;
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Dextranucrase Dsrd (EC 2.4.1.5).
 GN DSRD.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Neubauer H., Bauche A., Mollet B.;
 RT "Isolation and characterization of the dextranucrase Dsrd of
 RT Leuconostoc mesenteroides Lcc4.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY017384; AAG61158.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferrase; Glycosyltransferase.
 SQ SEQUENCE 1527 AA; 169835 MW; F9D0DE220BD89668 CRC64;

Query Match 80.4%; Score 86; DB 2; Length 1527;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRAHSEVQTVIAKII 22
 DB 654 NYSFVRAHSEVQTVIAQIV 673

Search completed: November 13, 2003, 09:44:09
 Job time : 35.1564 secs